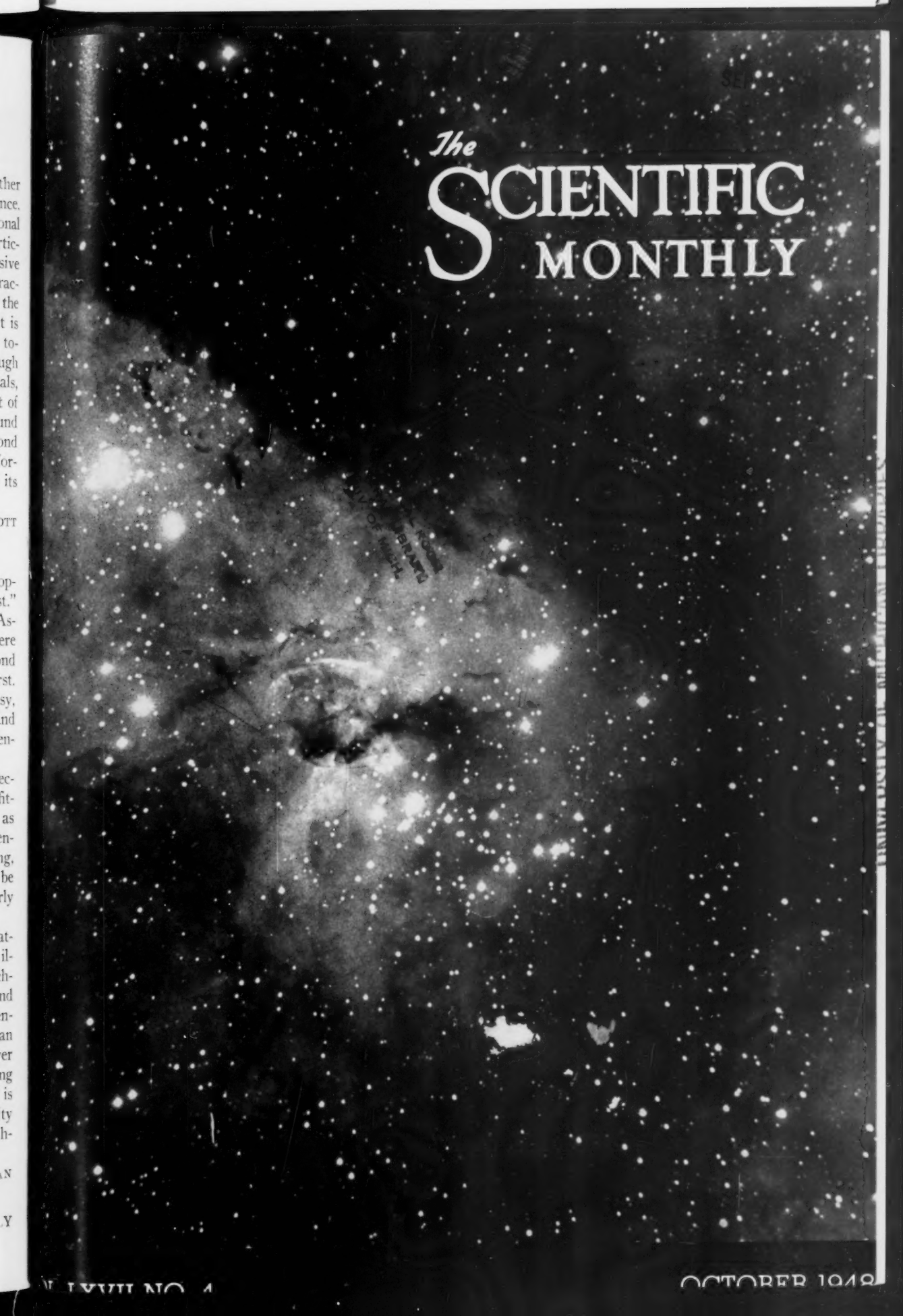
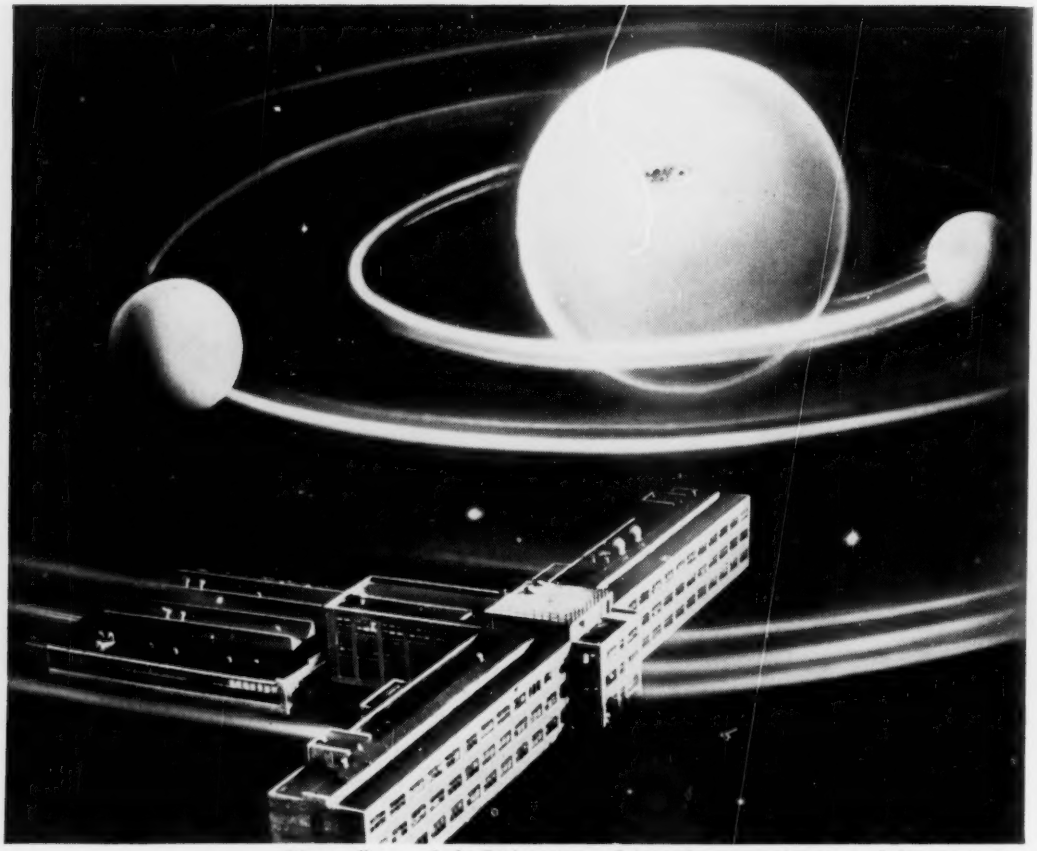


The SCIENTIFIC MONTHLY

VOL. LXVIII, NO. 4 OCTOBER 1949



"Sunspot" research, by RCA engineers, helps radio communications to dodge interference from magnetic storms. RCA Laboratories is a center of radio and electronic research.

93,000,000 miles of laboratory space

A cyclonic spot erupts on the face of the sun, and—here on earth—we feel it. Sunspots cause "magnetic storms," which disrupt radio communications.

What can be done about it? Research, during which RCA scientists and engineers "worked" by instrument on the face of the sun—93,000,000 miles away—offers an answer.

For many years, science has related magnetic storms to sunspots. An accurate way of forecasting these disturbances in advance was needed.

RCA scientists took a new tack. They noted that interference was most intense when sunspots were in a certain "critical area." *Location* and *activity* were more important than *size*.

Using this knowledge, RCA communications engineers now accurately forecast the beginning and end of magnetic storms. . . . They have established a daily magnetic storm forecasting service which is distributed like weather reports throughout the world. As a result, transmission of messages can be arranged over circuits or paths that will dodge interference.

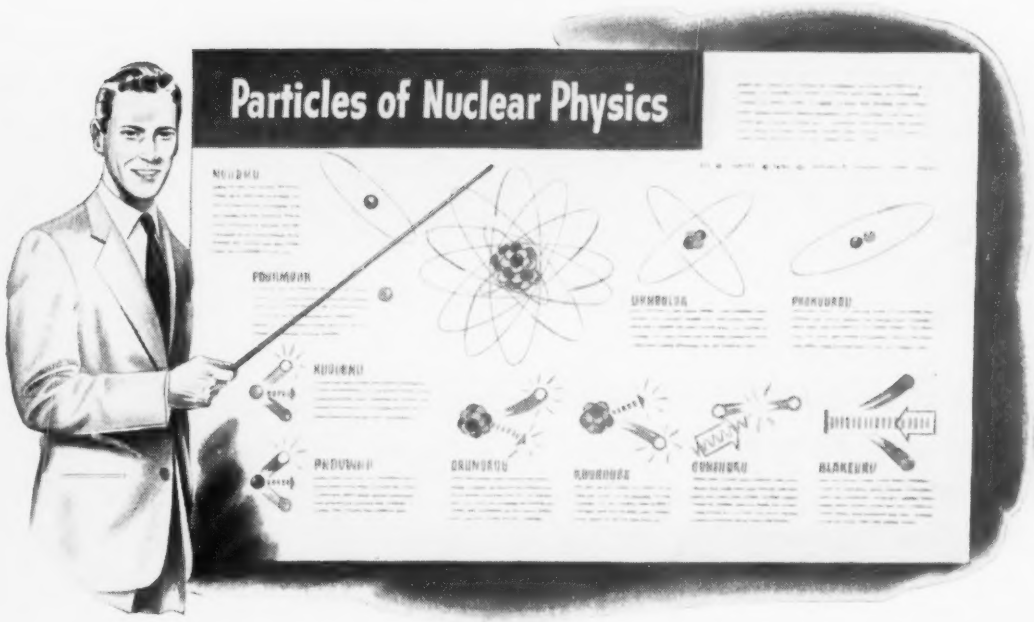
Such a pioneering spirit in research gives efficiency of service and leadership to all products and services bearing the names RCA, and RCA Victor.

• • •

When in Radio City, New York, you are cordially invited to see the radio, television and electronic wonders at RCA Exhibition Hall, 36 West 49th Street. Free admission. Radio Corporation of America, RCA Building, Radio City, New York 20.

RCA **RADIO CORPORATION of AMERICA**

YOU CAN BE **SURE** IF IT'S Westinghouse



Announcing 6 New Charts on
NUCLEAR PHYSICS

One of the most important offerings ever presented by Westinghouse School Service is this set of six wall charts on nuclear physics. To the teacher and student alike, they will prove invaluable in explaining this new science and in showing its significance in the world today.

The charts—lithographed in two colors—are a necessity in modern science classrooms. Included with each set is a 32-page booklet containing supplementary information on nuclear physics.

The six charts are:

1. ***Particles of Nuclear Physics:*** Portrays the ten basic particles important in the science of nuclear physics.
2. ***Structure of the Nucleus:*** Explains how nuclei are put together and charts 635 isotopes of the elements.
3. ***Nuclear Reactions:*** Depicts natural and man-made nuclear reactions, including nuclear fission.
4. ***Tools of the Nuclear Physicist:*** Portrays the important types of apparatus for detecting and inducing nuclear reactions.
5. ***Using Nuclear Energy:*** Shows five areas for the useful applications of atomic energy.
6. ***Progress of Nuclear Physics:*** Charts the major theoretical, experimental and engineering developments in nuclear physics. G-10036

School Service, Westinghouse Electric Corporation,
306 Fourth Ave., Box 1017, Pittsburgh 30, Pa.

Enclosed find $____ for ____ sets of Nuclear Physics Charts. (Price $1.00 per set.)

☐ Also send "Teaching Aids", listing 60 other free and low-cost classroom aids.

NAME_____
(Please type or print)

SCHOOL_____

ADDRESS_____

CITY_____ ZONE____ STATE____

Make check or money order payable to Westinghouse Electric Corporation.

Westinghouse
PLANTS IN 25 CITIES . . . OFFICES EVERYWHERE

THE SCIENTIFIC MONTHLY

VOL. LXVII OCTOBER 1948 NO. 4

Contents

Eta Carinae Nebula, Southern Milky Way (*John S. Paraskevopoulos*) . . Cover

Time and Change in the Metagalaxy . . . . . *Harlow Shapley* 243

Astronomer (*Verse*) . . . . . *Roland Ryder-Smith* 253

The Seven Degrees of Adventure . . . . . *Marston Bates* 254

Twentieth-Century Navigation . . . . . *N. C. Gerson* 257

Time (*Verse*) . . . . . *Francis Barry* 266

The Evolution of Terrestrial Life . . . . . *Addison Gulick* 267

Photography-in-Science Salon 273

A History of Population Growth in the United States . . . *P. K. Whelpton* 277

On Time as a Product of Motion . . . . . *Thomson King* 289

Three Methods of Voting . . . . . *William S. Taylor* 297

Science on the March:

Tropical Medicine—A Partial Report . . . . . *Eugene H. Payne* 302

Source Materials for Chemicals . . . . . *Gustav Egloff* 304

Book Reviews by *J. H. Bell, Th. Dobzhansky, A. G. Hall, Joel W. Hedgpeth,* and *Ralph F. Wolf* 308

Technological Notes 311

Comments and Criticisms 312

Published monthly at Business Press, Inc., by the American Association for the Advancement of Science under the direction of the Publications Committee: Farrington Daniels, J. E. Flynn, Kirtley F. Mather, Walter R. Miles, Malcolm H. Soule, and Steven M. Spencer. Gladys M. Keener, Assistant Editor. Publication office, 10 McGovern Ave., Lancaster, Pa. Orders for subscriptions and requests for change of address should be directed to the Circulation Department, A.A.A.S., 10 McGovern Ave., Lancaster, Pa., or 1515 Massachusetts Ave., N.W., Washington 5, D. C. Subscriptions: $7.50 per year; single copies 75 cents. Four weeks are required to effect changes of address.

Address all correspondence concerning editorial matters and advertising to The Scientific Monthly, 1515 Massachusetts Ave., N.W., Washington 5, D. C. The editors are not responsible for loss or injury of manuscripts and photographs while in their possession or in transit; all manuscripts should be accompanied by return postage. The American Association for the Advancement of Science assumes no responsibility for the statements and opinions expressed by contributors.

Copyright, 1948, by the American Association for the Advancement of Science. Entered as second-class matter at the Post Office at Lancaster, Pa., December 30, 1947, under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Sect. 538, P. L. and R., authorized December 30, 1947. Indexed in the *Reader's Guide to Periodical Literature.*

THE
SCIENTIFIC MONTHLY

OCTOBER 1948

TIME AND CHANGE IN THE METAGALAXY

HARLOW SHAPLEY

Dr. Shapley (Ph.D., Princeton, 1913) was astronomer at Mount Wilson Observatory during 1914–21, and since that time has been director of Harvard College Observatory. He has received numerous awards for his astronomical work, among them the Janssen Prix, Société Astronomique de France, 1933; Gold Medal, Royal Astronomical Society, 1934; and Pope Pius XI Prize, 1941. Dr. Shapley is the retiring president of the American Association for the Advancement of Science.

HOW MUCH of primeval chaos remains in the sidereal universe, and how much of order, is partly a question of definition. We are inclined, for working purposes, to define stars and their attendant planets, comets, and meteors as orderly parts of the cosmos, and leave in the chaotic category the interstellar and intergalactic dust and gas and the swirls of nebulous matter that are impelled through space by radiation pressure and other dissipative forces. But this definition is quite artificial. No true chaos exists for one who knows all the laws, and all the positions, motions, and masses of all material. In the microcosmic world, the Heisenberg principle would suggest that no such knower could exist. But we are speaking macrocosmically. We are not yet blocked by unknowableness. Our progress toward understanding the nature and the past and future history of the material universe will come through the formulation of partial laws, applying them to incompletely observed phenomena, testing the predictions that evolve from the assumptions and analyses, and thus amplifying the laws.

The Time here considered extends over a hundred million years; the Change cannot be directly observed. It must be inferred from contemporary evidence on objects near and far, and therefore on objects old and young, thanks to the perspective furnished by the finite velocity of light.

Fortunately for our progress, there are observable changes in the stars and even in the galaxies. We measure many kinds of variations: for example, the changes in position on the surface of the sky (very slight outside the solar system); and changes in motion, in light, surface temperature, and spectrum—but scarcely yet in age. From these measures, however, we try to infer the ages of various sidereal bodies, and also their chemical composition, dynamical history, direction of evolution, and destiny. The problems are large and our resources are small. The examination of the ratio of recognized order to apparent chaos, and the relation of change to time, are fields of inquiry that are discouraging to those who would avoid philosophical speculation and would stay near the observables. There seem to be too many vague puzzles. Still, when we note how great has been the progress in our understanding of the stellar universe during the past forty years, we infer that continued observation, and theoretical attacks on the problems we know enough to formulate, will probably provide during the next generation the grounds for demonstrating how primitive our present concepts are.

To help in making obsolete as many of our current views as possible, the Harvard Observatory has specialized in the study of the southern sky, which of necessity is somewhat neglected, because

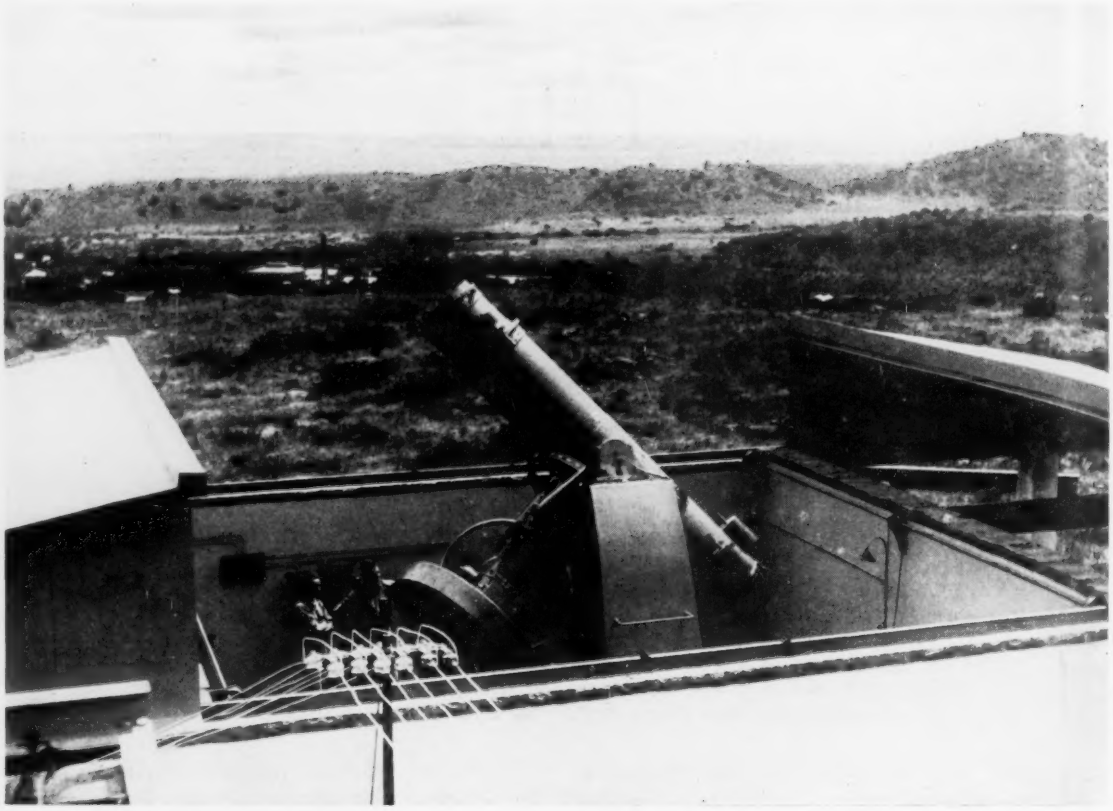


Fig. 1. 24-inch Photographic Refractor on Harvard Kopje, Bloemfontein, Orange Free State.

of the concentration of man power and mind power in the northern latitudes. Harvard's large southern station in the Orange Free State, South Africa, is dedicated in considerable part to studies of the metagalaxy—the over-all system of galactic systems. But the several researches reported in the present communication on "Time and Change in the Metagalaxy" will deal both with problems that could not have been handled from the northern observatories and with others that require work in the north only.

Concerning the metagalaxy. Although the metagalaxy was not in our science thirty years ago, we have now acquired a simply described picture of it. In a vast but perhaps not limitless expanse of space and time we find that the common luminous units are galaxies, of which the principal constituents are stars. The number of stars in a galaxy seems to be of the order of 10,000 million. But there is a wide dispersion in population. A supergiant galaxy, like our own Milky Way system, may contain more than 200,000 million stars, whereas some dwarf galaxies have scarcely more than 10,000,000. Indeed, there may be subdwarf galaxies of still smaller population; and already there is convincing evidence that some of the external star clouds, that we would now call galaxies, are probably no more massive, or luminous, or populous than the greatest of the near-by globular star clusters. Thus, the bright globular cluster 47 Tucanae, at a distance of only 25,000 light-years, is approximately as luminous intrinsically as the faint irregular galaxy IC 1613, at a distance of about 750,000 light-years.

Involved in the star fields of the individual galaxies, especially in those of spiral and irregular form, are both bright and dark nebulae, and groups of stars in all degrees of organization, from the compact million-starred globular systems and the clusters of the open Pleiades type to the multiple stars and simple binaries. Not so easily recognized as stars and nebulae is the widespread interstellar and perhaps intergalactic medium that is composed of electrons, atoms, molecules, and particles—the material (mostly nonluminous) that causes the absorption, blocking, and scattering of light.

This universal background stuff may be the most significant feature of the physical universe, with the stars only a phase that has monopolized human attention because they generate a radiation, as a

result of concentration of mass, that happens to affect strongly one human sense organ. Analogously, we have until recently considerable overrated the little segment of the spectrum from violet to red which our eyes recognize. To the "short" of violet and the "long" of red are the energies that empower and reveal the universe.

All these material entities, from interstellar atoms through stars to supergiant galaxies, are of interest in the study of the structure of the metagalaxy; but our attention will be given here mostly to the galaxies themselves, since they permit us to explore the depths of space more extensively than do the individual stars or the interstellar gases.

as-yet-unclassifiable, which is of course an enormously large group containing more than 90 percent of all now known. Distance, with consequent faintness and small dimensions on our photographic plates, hides from us the structures of most of the galaxies that appear on the long-exposure photographs made with the telescopes best suited to the recording of galaxies. The remote objects can be described only on the basis of shape and central concentration. Fortunately, however, a few thousand of the galaxies are near enough to permit somewhat detailed classification and analysis. We like to think that they are a fair sample of all the galaxies within 1,000 million light-years.

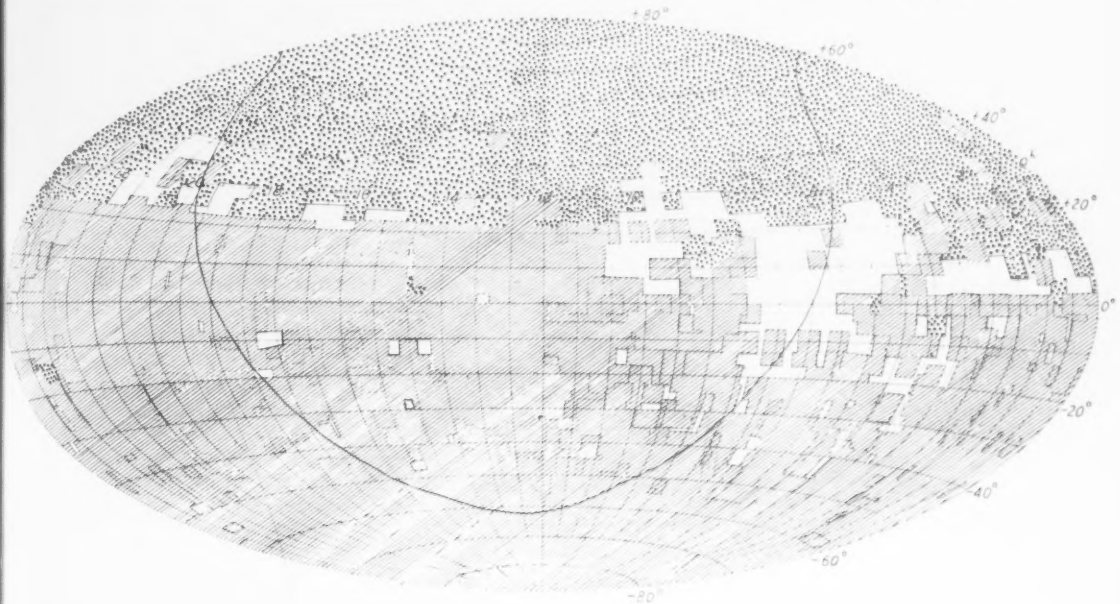


Fig. 2. Coverage of the sky in the eighteenth-magnitude metagalactic survey. Dotted areas were photographed with the Metcalf telescope.

The average separations of stars one from another are much greater, compared with their linear diameters, than the relative separations of galaxies. In ordinary regions of the metagalaxy there are less than 100 galaxy-diameters from one system to the next, and in some rich regions there is an average separation of not over 10 diameters; whereas from star to star in the neighborhood of the sun there is an average separating distance of more than 10,000,000 sun-diameters.

We find a considerable variety among the galaxies, especially in size and structural organization. But the inaccessibility of most galaxies, and our scanty knowledge of those within range, have led us to classify them into relatively few types: the spheroidal; the spiral; the irregular; and the

Our sketchy metagalactic picture can be concluded with a reference to the red shift in the spectra of external galaxies—a shift highly correlated with apparent brightness and therefore with distance, which has thus led to the hypothesis of the expanding universe. That the hypothesis is correct, and refers to all the space we survey, is probable, but not conclusively demonstrated. If we accept the hypothesis, we can refer to the age of the expanding universe. We find evidence that a few thousand million years ago there was an important epoch in cosmic history—an epoch that saw the beginning of the rapid expansion, the beginning also of the earth's crust (and of the earth itself), and a primitive arrangement of the star clusters in our own galactic system. How much of the evolu-

tion of stars and galaxies has occurred since that hypothetical zero point in time, we are not yet ready to say. Observations such as those to be summarized here should eventually contribute to our knowledge of the interrelation of Time and Change in the Metagalaxy.

Progress in a metagalactic survey. For a score of years one Harvard telescope in the Southern Hemisphere, the Bruce 24-inch refractor (Fig. 1), has been accumulating photographs, each of three hours' exposure on large fast-emulsion plates, as a part of a comprehensive survey designed to advance knowledge of the external galaxies in the southern and equatorial sky. Fifteen years ago a similar photographic doublet, the 16-inch Metcalf telescope, was moved from the Cambridge station to the new Oak Ridge station of the Harvard Observatory, and there we undertook to cover systematically the northern sky as a part of the same metagalactic survey. In Figure 2 we reproduce a map of the whole sky on an Aitoff projection—a projection which for most of the sky gives an equal number of square degrees for each square inch on the map. The curved line is the projection of the galactic circle.

The survey, as indicated by the figure, is more than seven eighths completed. The power of the plates varies somewhat over the sky because of atmospheric and seasonal variations, and because of inequalities from year to year in the speeds of photographic emulsions. But since plates of low quality are systematically retaken, we can claim a fair degree of uniformity for the survey. The aim is to photograph the stars of the whole sky to magnitude 18.0 ± 0.2; such photographs give in general a complete picture of the distribution of external galaxies to magnitude 17.5 ± 0.2. The corresponding distance in parsecs to which the survey reaches for an average galaxy in high galactic latitude, where space absorption is small, is

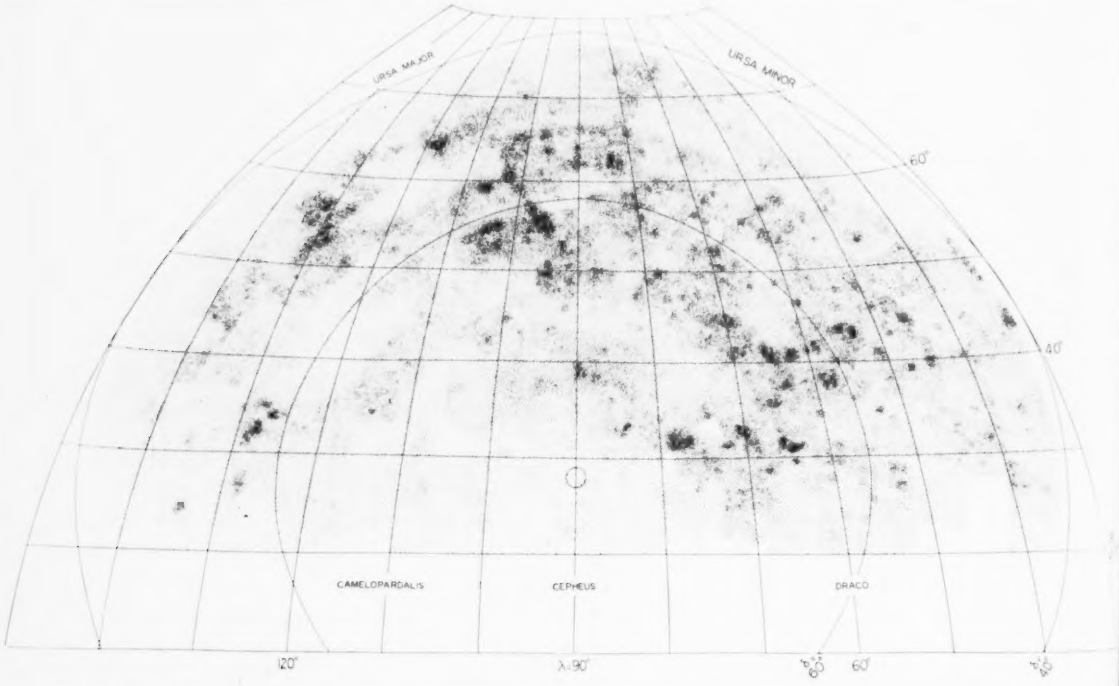


Fig. 3. The distribution of 80,000 galaxies in 3,600 square degrees north of Virgo.

computed from

$$\log d = 0.2(17.5 + 15.2) + 1$$

when, in a survey limited as this one is by apparent brightness, we take −15.2 as the average absolute photographic magnitude of a galaxy. The distance is 35 megaparsecs, or approximately 115,000,000 light-years.

The calculation refers to average systems only. The giant and supergiant galaxies, of course, appear in our census even when they are much more remote than 35 megaparsecs. Some as distant as 200,000,000 light-years are recorded. But occasionally dwarf galaxies at a distance of only 50,000,000 light-years are too faint to appear in our census. Most of the galaxies have luminosities distributed

so closely around the mean that we can fairly say that our survey covers metagalactic space to a distance of 100,000,000 light-years. In low latitudes, however, where the brightness has been dimmed by interstellar space absorption, the survey does not reach so far, and near the galactic circle it reaches nowhere at all because of the complete blocking by the clouds of obscuration along the Milky Way.

The long-exposure photographs already available in the eighteenth-magnitude survey show approximately half a million galaxies, more than nine tenths of which are recorded for the first time on these two series of plates. The completed survey, which will include many repeats of earlier photographs, should put on record, and into our statistical analyses, something more than 700,000 galaxies, of which perhaps 20 percent lie beyond the limit of 100,000,000 light-years. If it were not for low-latitude space absorption, the number within this distance could be nearly doubled. In other words, we find that at least a million galaxies are within 100,000,000 light-years of the earth.

If the mass of an average galaxy, including the interstellar material of its domain, is 10^{10} times that of the sun, the mean density in space is $\bar{\rho} = 5 \times 10^{-30}$ *gm/cc* throughout this surrounding volume of 4×10^{24} cubic light-years. That value is a lower limit. We have probably underestimated the number and mass contribution of the dwarf galaxies, which because of faintness will not be recorded in the Harvard survey; and possibly also we have underestimated the mass of interstellar material. Probably a better value is $\bar{\rho} = 10^{-28}$ *gm/cc*.

Out to the remotest distance at which an average galaxy could still be photographed with the most powerful reflecting telescopes, there should be at least 200,000,000 galaxies. That distance is approximately 600,000,000 light-years if space is effectively transparent in these outer regions. A supergiant galaxy like our own system could, however, be photographed with many of the telescopes now existing at double that distance. No doubt we have with the Bruce telescope already recorded objects a billion light-years away, but we have not identified the specific images of such galaxies among the thousands at the magnitude limits of our photographic plates. We deduce their presence statistically.

An atlas of northern galaxies. In the northern part of the constellation Virgo is a well-known cluster of some 200 bright galaxies, with magnitudes ranging mostly between 10.5 and 14.5. The diameter of the main body of the group, which is centered on R.A. $12^h\ 30^m$, Dec. $+12°$, is some 10°, but there are many outlying systems that, on the evidence of comparable brightness, could reasonably be assigned to this large cluster or cloud of galaxies. The most conspicuous extension of the cloud is to the south, through Virgo to the constellation Centaurus, at Dec. $-30°$. It makes the total extent about 5,000,000 light-years.

To the north of the Virgo cloud is a wide scattering of bright galaxies, covering the north galactic pole in Coma, and spreading over the constellations of Ursa Major and Canes Venatici. In our general study of this important northern region, we have photographed not only the bright galaxies, but also the background upon which they are superimposed. It is possible, therefore, to present, in the accompanying diagram (Fig. 3), a plot on an equal-area projection of the distribution of faint galaxies over all that area lying north of declination $+40°$ in higher galactic latitude than $+20°$. The survey plates were made with the Metcalf telescope at Oak Ridge, with three-hour exposures on fast emulsions.

The total area of approximately 3,600 square degrees covered by this survey contains about 80,000 galaxies, to the magnitude limit of 17.6, and many thousands fainter. Before making the diagram showing the distribution of the faint galaxies, all the plates have been reduced to a common magnitude limit. Appropriate allowance has been made for overlapping of plates, but no attempt to plot accurately the positions of individual galaxies within each square degree. Although in plotting we have somewhat exaggerated the smoothness of the distribution, the many real clusterings, deficiencies, and other irregularities are

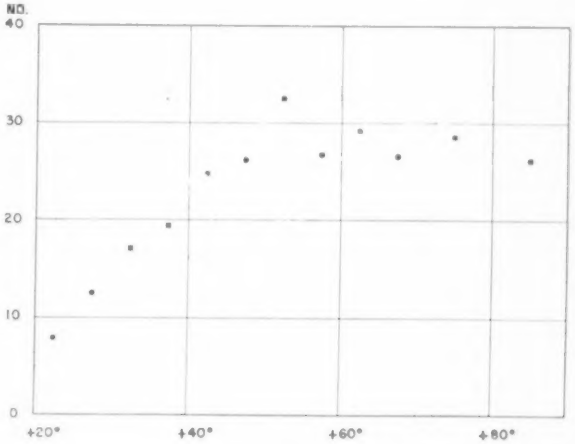


Fig. 4. Illustrating the effect of space absorption on average number of galaxies per square degree. Abscissae are galactic latitudes.

clearly shown. In particular, it is easy to see the effect of space absorption on the number of photographable galaxies in galactic latitudes from +20° to +30°. The conspicuous clusterings are now under special investigation with the aid of the Schmidt reflector at Oak Ridge.

A preliminary examination of the dependence of the number of galaxies in this part of the northern sky on the galactic latitude is assisted by Figure 4, where the average number of galaxies per square degree is plotted for 5-degree intervals of galactic latitude. (For latitudes greater than +75° the material is supplemented with observations not represented in Figure 3.) There appears to be no certain dependence on latitude above +40°. Below +40°, the number of galaxies per square degree falls off rapidly—more rapidly, in fact, than in most other parts of the sky. The strong space absorption shown here for latitudes between +20° and +40° in the northern sky is chiefly the "flare" of absorption in Cepheus, previously pointed out by Hubble,[1] and later confirmed by researches on the North Polar Cap by Shapley and Jones.[2] This Cepheid flare extends over the North Pole of the sky, and produces a considerable absorption and some reddening in the important field of the North Polar Magnitude Standards near Polaris.[2] The flare is limited in galactic longitude, as can be seen in Figure 3.

It is of passing interest to note that we have photographed 1,760 galaxies through the bowl of the Big Dipper. The galactic coordinates of its center are $\lambda = 105°$, $\beta = +58°$.*

The galactic anticenter. Diametrically opposite the galactic nucleus, which lies chiefly in Ophiuchus, Scorpio, and Sagittarius, is a less significant but nevertheless important region of the sky. In the anticenter region we can hope to discover most easily the extent of the Milky Way in its own plane and effectively explore the outer regions of a spiral galaxy.

Obscuring matter along the galactic circle in Taurus and Auriga, the anticenter constellations, blocks measurements in low latitudes. Any useful distance indicator requires transparency, or quantitative knowledge of the absorption. By making a careful simultaneous study of Cepheid variable stars and external galaxies in latitudes higher than 15 or 20 degrees, on both sides of the Milky Way, we can, however, get at least an approximate measure of galactic extent in the direction opposite the center. The periods and brightness of the Cepheids along the borders of the Milky Way give us the first estimates of distances. But they give maximum values. The number of faint galaxies provide a measure of the total amount of space absorption in each field photographed for variables and, therefore, indicate how much correction must be made to the photometrically determined distances of the Cepheid variable stars in order to get reliable values.

The systematic study of the anticenter variables has been in progress for five years. The region under survey is shown, in galactic coordinates, in Figure 5. Several hundred plates have been made for the discovery and measurement of the faint Cepheids. Various photographic telescopes are being used, chiefly those located at the Oak Ridge station; the most effective for the work, because of large field and deep reach, are the 16-inch Metcalf refractor and the 24-inch Jewett Schmidt-type re-

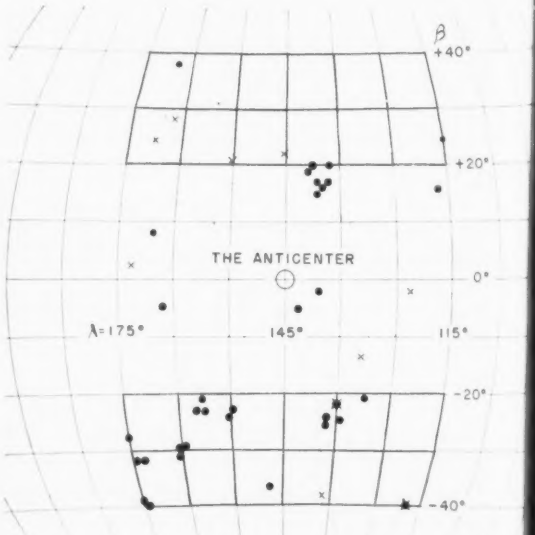


Fig. 5. Diagram showing the two anticenter regions under special exploration for Cepheid variable stars. The cluster variables in this part of the sky that were known at the time work was begun are shown as dots when brighter than photographic magnitude 12, as crosses when fainter. Those close to the galactic circle, $\beta = 0°$, are generally too much troubled by space absorption to be useful. More than 100 new variables have now been discovered in the two regions, and many are cluster-type Cepheids.

* The examination of the photographs for the work described in this section has been chiefly in the hands of Rebecca Jones, with assistance from Frances Wright and others. The photographs with the Metcalf refractor were nearly all made by Henry Sawyer at the Oak Ridge station. (Detailed description of the work will be published elsewhere.) The dependence of frequency of galaxies on galactic latitude can be better analyzed after the photographic magnitudes of the individual objects have been determined.

flector. The variables previously known and catalogued in the anticenter region are in general too bright and too near, or too much involved in the lower-latitude absorbing clouds, to be useful in this work. The 150 variables, newly discovered for this study by Martha Dowse, are mostly fainter than the fourteenth magnitude, and many are located in the regions where the population of galaxies is sufficient to indicate high space transparency.

Of the new variables, many are of the cluster type, and at this stage of the research, we can report that several are as far from us in the anticenter direction as we are distant from the galactic center. Presumably these faint variables are members of the galactic system, not intergalactic; but some of them may lie in the "star haze" surrounding the discoid of the galaxy,[3] rather than in the main body of our flattened spiral system.

Eventually the research on the anticenter region will involve star-counting and the analysis of the colors and spectra of faint stars, as well as the present exploration by way of galaxies and Cepheid variables.

Variable stars in the Clouds of Magellan. The Magellanic Clouds (Fig. 6), as the nearest of external galaxies, are important in our study of the metagalaxy because of their resolvability, their irregular forms, and their significance in the local cluster of galaxies. Their Cepheids are of wide cosmic usefulness. They provide much information valuable to the study of Cepheids in our own galaxy. Recent investigations in the Clouds have brought out some new facts about Cepheids and have reinforced others already surmised. A few may be summarized as follows:

a) Continued study of the photographic plates has increased in the two Clouds the number of recognized classical Cepheids (periods greater than a day) to more than 2,500. Some scores of irregular variables and eclipsing stars, found and studied on these photographs, are mostly giant and supergiant members of the Clouds, rather than intervening galactic variables. Many of them are a thousand times the luminosity of the sun.[4]

b) About a dozen cluster-type Cepheid variables (periods less than a day) are found on the photographs of the Small Magellanic Cloud, but probably without exception these objects are all superposed. No cluster variables that are definitely members of either Magellanic Cloud have been found, notwithstanding the use of a special series of plates made with the 60-inch reflector at the South African station. Certainly such stars are

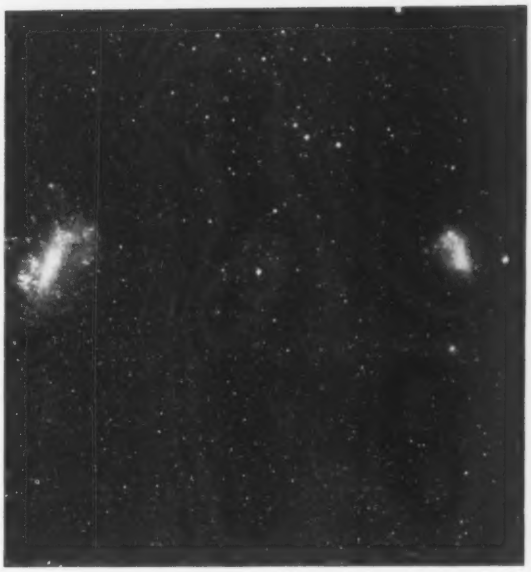


Fig. 6. Magellanic Clouds, the nearest external galaxies. The globular cluster 47 Tucanae (Fig. 10) appears at the edge of figure to right of the Small Cloud.

very rare in the Clouds, in striking contrast with the situation in the galactic system where the cluster-type Cepheids are much more abundant than the classical Cepheids.

c) In the core of the Small Magellanic Cloud are many classical Cepheids with periods abnormally long—10–20 days—with scarcely any with the dominant period length, that is, of 2–5 days. A similar scarcity is shown in our accumulating data on the globular star clusters, and it is particularly conspicuous for classical Cepheids in the direction of the galactic center, where the "16-day Cepheid" is especially dominant. In all three regions[5] the star density is high. This peculiar distribution of the periods is doubtless a matter of evolutionary significance in the history of rich stellar systems.

d) Certain regions of the Small Cloud have been thoroughly studied for the magnitudes and periods of the variables. The danger of preferential selection, on the basis of period or brightness, is thus eliminated. The distribution of the lengths of the periods in these thoroughly explored regions is shown in Figure 7. The maximum frequency appears for periods of 2.5 days, whereas, for the regions in the solar neighborhood, the maximum is between 4 and 5 days. The contrast is in small part a matter of incompleteness of the surveys in our own galaxy; but clearly it indicates a basic difference in Cepheid phenomena in adjacent galaxies.

e) As a further result of the study of the periods

and light curves of more than 500 of the Cepheid variables distributed all over the Small Cloud, we have confirmed, as noted above, the fact that Cepheids of longest periods show a preference for the dense central regions, and that the shortest periods predominate in the intermediate and boundary regions. This peculiar phenomenon has not yet been established for the Large Magellanic

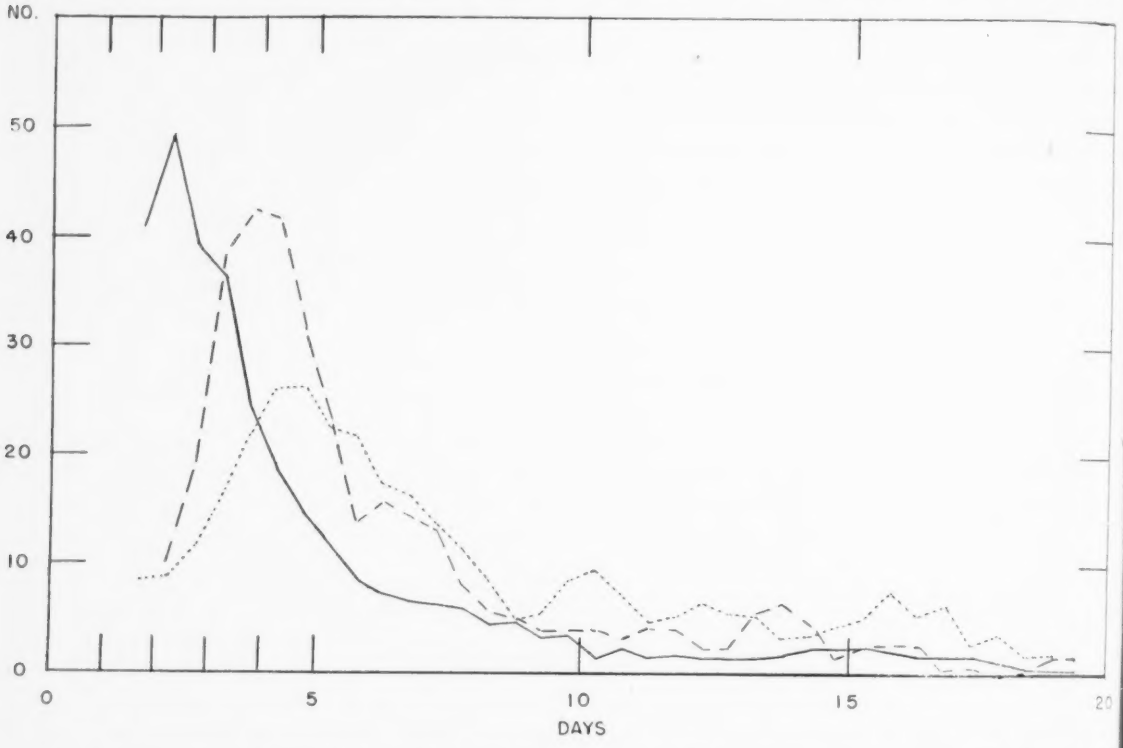


Fig. 7. The frequency of periods of Cepheids. *Full line*: 531 Small Cloud variables; *broken line*: 144 Large Cloud variables; *dotted line*: 288 galactic variables. For convenience of comparison, the curves have been reduced to a common total of 250 by multiplying the ordinate for each plotted point by 0.47, 1.74, and 0.87, respectively.

Cloud where, so far, much less material is available and where, very surprisingly, scarcely a variable with period less than 2.5 days has been found.

f) A detailed analysis of the dispersion of magnitudes about the period-luminosity curve shows that half a dozen factors contribute to the spread in brightness. The four most important are the intrinsic differences in luminosity for stars of a given period; photometric measuring errors, including the doubling effect in the crowded star fields; thickness of the Cloud in the line of sight; and patches of absorbing material within the Clouds.

g) An apparent extension, or wing, of the Small Magellanic Cloud was recently discovered. Its distance has now been determined by means of Cepheid variable stars, and it is shown to be a physical part of the Small Cloud—not an independent system at some other distance. The wing extends toward the Large Cloud.†

Distribution of luminosity in galaxies. With the assistance of Dr. Patterson, I have begun systematic studies of the distribution of luminosity over the surface of the nearer spheroidal and spiral galaxies. The photographic work has been done at the Oak Ridge station with a flat-field refractor, and the plates have been carefully standardized throughout the processes of exposing, developing, calibrating with the sensitometer, and analyzing with the microdensitometer. Several workable galaxies generally appear on each plate. About half the photographs have been reduced and discussed, and a report on one phase of the work has already been published. Another discussion will appear in

† In the investigations of the Cepheid variable stars in the Magellanic Clouds, reported in the foregoing summarizing paragraphs, I have been assisted by Virginia McKibben Nail, Richard A. Craig, Frances W. Wright, and Martha Dowse. The photographic work in the Southern Hemisphere has been under the direction of John S. Paraskevopoulos, superintendent of Harvard's South African station.

a Harvard Observatory publication. Two of the principal results may be summarized:

a) The measurements of a number of spiral galaxies, whose galactic planes are highly inclined to the line of sight, indicate that only about 20 percent of the light, sometimes much less, of a typical spiral is located in the spiral arms themselves. Some of it is in the nuclei, of course, but most of it is in the little-noticed background of the galaxy. The evolutionary implications are obvious. *b*) The over-all dimensions of galaxies of the various types appear to be much alike. Earlier studies had seemed to indicate that the open spirals were several times more extended than the spheroidal galaxies. But it now appears from the microdensitometry that the spheroidal structureless systems can be traced to great distances from their nuclei. The arms of a spiral galaxy (and three fourths of the galaxies show these spiral arms) appear, therefore, to be concentrations, or perhaps condensations, within the systems rather than ejections from a central nucleus. In Figure 8 the distribution of the diameters of the various types of galaxies, as shown by the new series of measures, is illustrated, with the abscissae expressed in kiloparsecs.[6] The measured diameters of the smallest objects are about 6,000 light-years, the median value is twice as great, and the largest systems are more than 30,000 light-years in diameter and therefore approach comparability with our own giant galaxy.

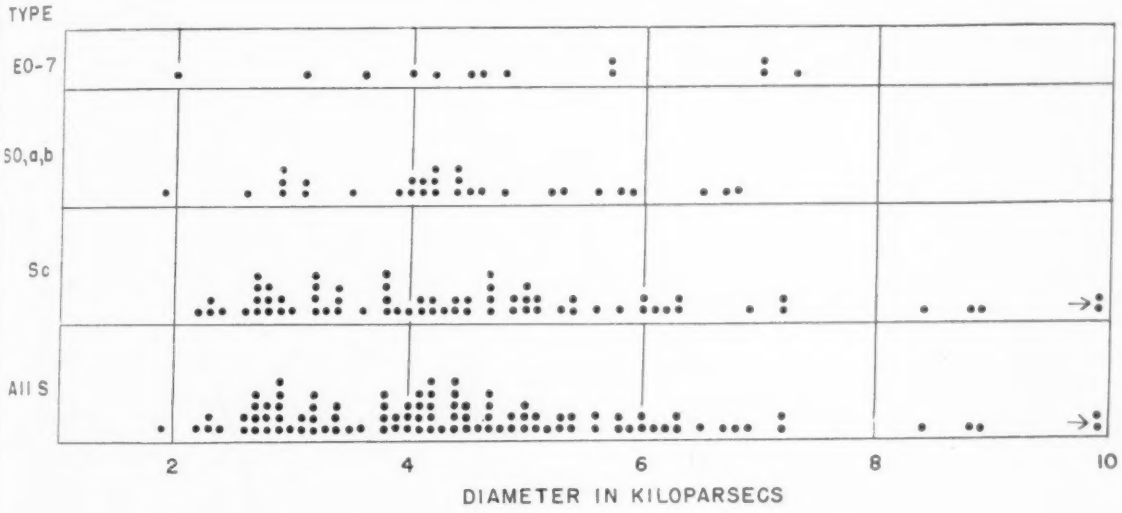


Fig. 8. Distribution of dimensions of galaxies. One kiloparsec equals 2×10^{16} kilometers.

On the evolution of galaxies. From the foregoing new observations on the dimensions of galaxies and the significance of spiral arms, one might provisionally conclude that the open spirals do not necessarily represent the older or more developed state of a galactic system. If there is development along the well-known sequence of galaxy forms, it may be in the direction opposed to that usually assumed; that is, in the direction from the most open spiral (Sc, Fig. 9) to the less open spirals (Sa) to highly flattened spheroids (E7), toward the spherical galaxy (E0). In such a scheme globular clusters like the giant 47 Tucanae (Fig. 10), or the still more gigantic globular Messier 87, which is a typical spheroidal galaxy, might be considered the end products of galactic evolution. One rather colorless argument for this hypothesis is the above-mentioned suggestion that spiral arms may not represent rotational ejection from the nucleus of a developing galaxy, but represent instead merely an inner organization of the nebulosity, star clusters, giant and supergiant stars. But such an interpretation of the arms does not necessarily define the direction of development. Positive arguments for a spiral-to-spheroidal trend in galactic evolution are the following:

a) Because of the shearing action in the rotation of a galaxy of Class Sb or Sc (our own galactic system, for instance), the star clusters and star clouds which characterize these open galaxies will be gradually dismembered and the whole system must proceed toward structural uniformity. Thus, as Bok and others have shown, the trend of loose star clusters is toward dissolution. We have visualized no reasonable mechanism for constructing star clouds or loose clusters of giant and supergiant stars out of the uniform star fields that characterize the structureless spheroidal galaxy. If evolution were in the direction of spheroidal to spiral, what could be the mechanism that develops irregular clustering of giant and supergiant stars?

b) The radiant life of supergiant stars is relatively brief, according to current ideas on the

atomic generation of stellar energy. Spitzer and Whipple have discussed a process by which the red supergiant stars can develop out of dust clouds. But such stars apparently do not and cannot evolve in a spheroidal galaxy; my observations on the giant globular clusters indicate that interstellar dust does not exist profusely if at all in a globular cluster or in a spheroidal galaxy, which is therefore a sterile place for the production of giant young stars. Open spirals and irregular galaxies, on the other hand, are characteristically full of dark and bright nebulosity. There the material is available for the birth of giant stars of low density.

Fig. 9. Messier 83, a bright southern spiral of class Sc.

The time required for the transition from a very open spiral, with its star clusters, nebulosity, dust clouds, and supergiant stars, to a globular galaxy from which all this population, as well as the flatness and spiral structure, has disappeared, would undoubtedly be so long as to strain severely the short time scale of about 10^{10} years. The short-time scale, however, should not be held as conclusively demonstrated, notwithstanding strong evidence favoring it, such as the age of the earth's crust, the speed of the expansion of the universe, and the present existence of open star clusters in the Milky Way. But even within the framework of this limited time schedule one could look at the development of galaxies from the standpoint of different speeds in two different phases; the first being the rapid and explosive adjustment, when the universe was "young"—an adjustment which resulted in sidereal matter aggregating into unit galaxies of

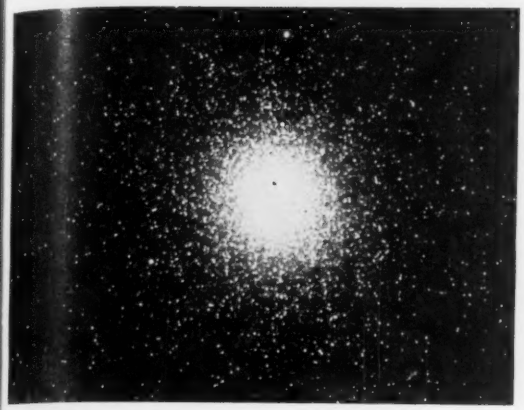


Fig. 10. The globular cluster 47 Tucanae.

many sizes and forms, much as we see them today; and the other phase being that more deliberate dynamical and radiational process that is now going on, which in the long run must tend to smooth out and perhaps round up both the irregular galaxies and the considerably nucleated open spirals. There is much room here for theoretical analysis.

Of the five brightest systems in our local group of galaxies, two are irregular (the Magellanic Clouds); of the total group of a dozen, half are irregular—a much higher proportion of chaos-touched systems than we observe elsewhere in metagalactic space. If these irregular star clouds were isolated from perturbing giant galaxies like the Andromeda Nebula and our own system, they might eventually become much smoother. The Andromeda Nebula and our own galaxy both have faint approaches to sphericity in the apparently spheroidal haze of distant high-latitude stars, which for our system includes the high-velocity "escape" stars of the cluster-variable type. But are these spheroidal hazes, which surround the massive discoids, incipient or vestigial?

REFERENCES

1. Hubble, E. The Distribution of Extra-Galactic Nebulae. *Astrophys. J.*, 1934, **79**, 35; *Mt. Wilson Observatory Contribution 485*, 1934.
2. Shapley, H., and Jones, R. B. Survey of 16,639 Galaxies North of Declination +70°. *Harvard Ann.*, 1938, **106** (1), 1–361.
3. *Harvard Reprint 173; Proc. Nat. Acad. Sci.*, 1939, 25, 423–8.
4. *Harvard Reprint 306*, 1948; *Proc. Nat. Acad. Sci.* 1948, **34**, 178.
5. *Harvard Reprint 192*, 1940; *Proc. Nat. Acad. Sci.*, 1940, **26**, 105; *Harvard Reprint 241*, 1942; *Proc. Nat. Acad. Sci.*, 1942, **28**, 200.
6. *Harvard Reprint 238*, 1942; *Proc. Nat. Acad. Sci.*, 1940, **28**, 189.

ASTRONOMER

All night he watches roving worlds go by
Through tempered glass, his window on the sky,
Feels in his own the beat
Of some far mightier heart, and hears
The mystic concert of the spheres.

Scanning infinitude, that dim vault kissed
By light, he centers on one star long missed.
For this unstable ray,
This bright but strangely prodigal beam,
The torch within his mind will gleam.

Roland Ryder-Smith

THE SEVEN DEGREES OF ADVENTURE

MARSTON BATES

Dr. Bates, an entomologist (Ph.D., Harvard, 1934), has been associated with the International Health Division of the Rockefeller Foundation since 1935, working in Albania until 1939, and then spending a year in Egypt. In 1940 he moved to Villavicencio, Colombia, where he is working on jungle yellow fever.

SCIENTISTS frequently become involved in the problem of writing books about their travels in the tropics. The question of adventures, in such circumstances, must be taken very seriously. It may be possible to explore the rain forest without encountering adventures, but the moment you use the word "jungle," the reader starts looking for an adventure, and he will be happier if adventures occur at fairly regular intervals throughout the text. Keeping the reader happy is an important consideration; if he gets bored, he may shut up the book and go off to a movie, or even remember that he is supposed to help his wife wipe the dishes.

I started to write a book about the tropics the other day, and these factors led me, logically, to undertake an analysis of the phenomena of adventures. Clearly, adventures are natural phenomena and, as such, subject to attack by the ordinary methods of scientific analysis. Preliminary study of this sort should enable any scientist to handle his adventure problems in a reasonable way, and the results of my study may be helpful to colleagues who find themselves in a similar situation.

I have always entered adventures in my journals, along with the more routine natural-history observations; in addition, I have had the foresight to make notes on the adventures of friends and casual acquaintances. These made a rather unwieldy accumulation that plainly would have to be reduced to some sort of order before it would be possible to compose any coherent narrative. After various preliminary tests, I decided that a punched-card system would provide the best method of analysis.

Punched cards are an indispensable tool in scientific work. I don't mean the punched cards used by the statistical Johnnies with big budgets: cards with orderly patterns of holes punched by electric machines, those superhuman monsters that sort the cards, count them, digest them, and, finally, with perspicacious efficiency, flush them down the drain. The common, nonstatistical, scientist uses another kind of punched card. It has rows of holes all around the margin and ample space for writing in the middle. You write out the data (adventure, in this case) on the card, and then study its characteristics. You cut notches into different holes along the margin, according to these characteristics, and then file the cards.

If, for instance, the adventure includes a poisonous snake, you convert the third hole from the left on the top of the card into a notch; if the adventure concerns a drunken Indian trying to stab someone in the back, the fourth hole from the left, and so on. Thus, if you want to compare all the adventures involving drunken Indians with knives, you simply borrow your wife's knitting needle, run it down the pile of cards through the fourth hole from the left on top, shake the cards, and all those with drunken Indians fall out.

I must acknowledge that my secretary, Miss Carmichael, did most of the work. I went through my journals carefully and marked my adventures with a red pencil; she then copied each one out on a punched card. We worked out the analyses together. We soon found that there weren't enough adventures in my journals for statistical significance, so Miss Carmichael undertook a general review of books on travel in the tropics, making cards for all the adventures. This kept her busy for a long time and helped me a great deal, since I didn't have to think up things for her to do in the mornings.

We found that adventures were most satisfactorily classified according to a scale of degree of intensity, and that they fell into seven categories. Adventures of the First Degree are those in which the author faces death. Unfortunately, I never knowingly faced death in the tropical forest, so there were no examples for study in my journals. We found many fine ones, however, in the travel literature: the subconscious "freeze" just as the intrepid explorer is about to tread on a bushmaster (*Lachesis mutus*); the drunken native lunging in the dark with his machete; the rifle that misfires as the cornered lion springs forward in fury.

Miss Carmichael collected two hundred and six

Reprinted by permission from *The New York Times Book Review*

"What a footnote this'll make to Chapter Five!"

en specimens of Adventure of the First Degree m her meticulous study of seventy-six volumes f travel and exploration in tropical regions. From hese we defined certain peculiarities of the circumstances under which the authors of these volmes faced death. There were poisonous snakes in 9.2 percent of the adventures; in 5.1 percent of hese, however, we were able to determine, from ndirect evidence inadvertently included by the uthor, that the snake was really of a harmless ind. Authors should be more careful about these hings. If they expect to face death from a poisonus snake in the water, they should first make sure hat there are poisonous water snakes in the region hey are exploring.

Hostile Indians, drunken savages, belligerent atives, or other varieties of *Homo sapiens* were nvolved in 13.5 percent of the adventures. The uthor always gave details from his point of view, at in no instance was there any adequate statement of what provocation he might have given hese Indians, savages, or natives.

Mammals other than man constituted the peril in 26.3 percent of the cases. There was a nice geographical correlation here, since the majority of these adventures occurred in the Old World tropics. Even a seasoned explorer has trouble achieving First Degree Adventure with the tropical American mammals.

It is remarkable that the author was not actually killed in any of the two hundred and sixteen situations. (Of course, if he had been killed, he wouldn't have been able to finish the book, but you would think that maybe his wife could have taken over or something.) We got interested in this and checked on the longevity statistics for tropical explorers. It turned out that on the average they live to a very ripe old age, generally dying of pneumonia in England or in an automobile accident in the United States. I feel strongly about this because, although I never experienced any First Degree Adventures in the tropics, I have frequently faced death in Harvard Square or at

Broadway and Forty-second Street. Perhaps I could work that into my book by including a final chapter on Facing Death in Harvard Square (although it looked for a while this morning as though my wife might have to write it as an epilogue, for posthumous publication).

Next there are Adventures of the Second Degree. These have to do with women. Miss Carmichael looked carefully through my journals, but was unable to find any. (My wife also has access to my journals, so this is perhaps not strange.) It is queer, though, that she found no examples in the seventy-six volumes of travel and exploration, despite the fact that slightly more than half these books were written by bachelors. Surely they must have been exposed to Adventure of the Second Degree at times on their travels: a willowy blond from Texas in the cocktail lounge of the steamer; a shy, dusky maiden bathing in a lonely stream in the forest; some unforgettable figure in the wild abandon of a savage dance propitiating some god of fertility. By all logic of human nature, these authors must have encountered Adventure of the Second Degree, but they are remarkably reticent about it.

Adventures of the Third Degree involve sublethal danger; of the Fourth Degree, pests; and of the Fifth Degree, officials. These are all very interesting categories. Sublethal danger is a particularly protean phenomenon: it includes disease (fatal diseases, of course, are never listed among adventures); hunger; being lost in the woods; bodily injuries and mutilations; and major inconvenience like missing airplane connections.

Pests (Fourth Degree Adventures) include perhaps the commonest and most-written-about incidents of life in the tropics. The catalogue is endless: scorpions, mosquitoes, black flies, fleas, bedbugs, centipedes, redbugs, ants, and my wife's friends, the cockroaches. It is perhaps arbitrary to separate officials (Fifth Degree Adventures) from other kinds of pests, but the separation has convenience, since adventures with officials and with insects generally take place in different settings. Official adventures in themselves may have considerable variety: customs officers puzzled by pickled snakes; police incensed by the photographs of strategic defenses such as colonial fortresses; poker games with *El Presidente* (travelers always seem able to contrive the most friendly relations with presidents and kings); getting a visa when you have forgotten to bring along three photostatic copies of your marriage certificate.

There remain only the Absurd Adventures (Sixth Degree) and the Unforgettable Thrills (Seventh Degree). In classifying absurd adventures Miss Carmichael always tactfully took the author at his face value. If he reported ordering a glass of milk and getting a fried chicken instead, that was an absurd adventure, introduced to add comic relief. If he called to his gunbearer for the high-powered rifle to shoot a rhinoceros, and was given a shotgun instead, that was sublethal danger (he could always club the rhinoceros with the butt of the shotgun if it attacked). The first example, of course, illustrates the author's struggles to learn the outlandish native language; the second illustrates native stupidity and unreliability in time of stress.

Unforgettable Thrills (Seventh Degree) are always easily recognized: "a breath-taking panorama was suddenly spread before my eyes;" "no white man had ever before heard the clear, lilting song of this beautiful bird;" "there, quivering in the palm of my hand, lay a new butterfly, unknown to the catalogues of science;" "the sheer beauty of that flaming orchid will forever be etched in my memory." It may be significant that Miss Carmichael found a slightly higher frequency of unforgettable thrills among male than among female authors.

When we started to write my book, "Wanderings in the Jungles of South America," we found that Miss Carmichael had mixed up my adventures with those culled from the seventy-six previous volumes of travel and had, unaccountably, thrown them all away. There was enough Whole Cloth around the office, however, so that we were able to proceed with the preparation of the manuscript.

TWENTIETH-CENTURY NAVIGATION

N. C. GERSON

After graduating from the University of Puerto Rico in 1943, Mr. Gerson, whose special field is meteorology, entered the Technical Investigations Section of the U. S. Weather Bureau. He later became associated with the Low Frequency Loran Experimental Project in southwestern Canada. At present he is chief of the Electromagnetic Wave Propagation Section, Watson Laboratories, Red Bank, New Jersey.

Now there be some that are very inquisitive to have a way to get the longitude but that is too tedious for seamen since it requireth the deep knowledge of astronomy, wherefore I would not have any man think that the longitude is to be found at sea by any instrument; so let no seamen trouble themselves with any such rule, but (according to their accustomed manner) let them keep a perfect account and reckoning of the way of their ship.—Nautical works of seventeenth century.

IN ITS general sense, navigation may be defined as the science of location, i.e., the science by means of which one determines his position within a given coordinate system from observations made upon objects either within or without the system. From a knowledge of his position, the navigator of a mobile unit may then set a course to any other object whose coordinates are known. Thus, upon our planet, the science of navigation affords the perception necessary to conduct a vessel or other surface craft from point to point on the earth's surface, and enables the navigator to determine his position at any time within some minimum degree of accuracy.

Although in present usage the connotation of navigation has been expanded to include avigation, nevertheless some important differences between the two sciences should be recognized. Navigation is restricted to surface location; i.e., it pertains basically to a surface craft plowing through the water or moving over land at a low velocity—known to within a fairly close limit. With a vessel, ocean currents, winds, tidal action, etc. may force the ship from its course. However, because most of these deviating factors may be estimated accurately, corrective action which practically eliminates their effect may be taken. Contrast this somewhat simple practice with that normally found in avigation. Avigation, restricted to location above the surface of the earth but in the atmosphere, deals fundamentally with aircraft traveling at a high rate of speed. The velocity, known only approximately, takes place in a supporting medium which usually is in a highly variable and complicated state of motion. As atmospheric winds cannot be plotted in the manner of hydrospheric currents, since the operation occurs in three dimensions instead of two and because aircraft travel rapidly rather than slowly, avigation is generally more difficult and more hazardous than navigation.

Notwithstanding the finer distinctions between navigation and avigation, the present discussion will conform with popular usage and allow the former term to connote also the latter. Thus, the original definition of navigation may be expanded to include location either on or above a set of coordinates from observations made on objects, whose coordinates are known, residing either inside or outside the system or both.

It will be interesting to note the various subdivisions of navigation as the science is applied on the earth's surface. All navigation by observations made upon objects within the coordinate frame of meridians and parallels is lumped together under "geonavigation," whereas when the objects are extraterrestrial, the science is termed "celestial navigation." When it is necessary to resort to extrapolation procedures, the method is termed "dead reckoning." A convenient breakdown is shown in Table 1.

TABLE 1

NAVIGATION

Geonavigation	Pilotage
	Sonic methods
	Radiolocation
Celestial navigation	
Dead reckoning	

Pilotage is analogous to motoring or perhaps walking. Both pilot and pedestrian move from origin to destination by noting visually the location of prominent landmarks along the route. Mountains, towns, and rivers are significant in the case of the pilot, and buildings, street signs, and bridges in the case of the stroller. Except under special circumstances, pilotage is invariably employed when possible, even though other navigational aids are present. Obviously, pilotage is impossible in fog, clouds, over oceans, or whenever visible landmarks are absent.

Sonic types of navigation depend upon the re-

flection of compressional waves traveling either in air or in water. The time required for a whistled signal or other sound to travel from a vessel to a near-by cliff and return is measured. From the known speed of the wave, the distance between the ship and cliff is quickly determined. Similarly, reflections of sonic waves from the ocean floor may be utilized. Directional receivers may be employed similar to electronic DF systems. Present sonic methods are confined to vessels, not aircraft. With the depth contours now appearing on the nautical charts published by our Federal mapping agencies, the navigator, equipped with an echo-sounding device, can use the many submarine features delineated on the charts for position location.

Radiolocation, the third division of geonavigation, is coming into widespread usage. In the early future it may become the principal mode of terrestrial navigation, especially if combined with some form of an automatic dead-reckoning device. Radiolocation employs the radio wave region of the electromagnetic spectrum as an aid in position determination. Although its prewar usage was mainly confined to such devices and systems as radio compasses, radio range stations, absolute altimeters, etc., today it embraces a wide variety of competing electronic orders, most of them hidden under odd or fanciful names. Some of these systems are known as Gee, Shoran, Racon, Consol, Decca, Raydist, Loran, etc. (It should be noted that if instead of employing radiolocation as a category under geonavigation, *electromagnetic location methods* had been used, then pilotage, which employs the visible region of the spectrum with the eye as a receiver, also would have been included within this classification.)

Celestial navigation is self-explanatory. Its practice consists of establishing the position of at least two heavenly bodies upon the celestial sphere from the observer's site. By a transformation of coordinates the observer's position upon the earth is then determined. Celestial navigation is the most complex of the various forms of navigation and is usually employed over the oceans, over cloud banks, over isolated tracts of land, etc., when pilotage or radiolocation is not available or when an independent position check is desired.

Dead reckoning has been the major form of navigation thus far utilized. Essentially it consists of maintaining a record of the velocities and times which have elapsed since leaving the last positively identified position. Dead reckoning, imperative when pilotage is impossible, suffers one great disadvantage. Since errors in it are cumulative, the use of dead reckoning is limited by the unknown factors. An unknown wind velocity effective on a long course, for example, may carry a pilot far from his final destination. Between position determinations obtained by means of either geonavigation or celestial navigation dead reckoning, of course, is used.

With the exception of radiolocation, one major interfering factor is common to all methods of navigation employed below an altitude of about 25 km—weather. Rain, snow, fog, dust, or other meteorological conditions attended by poor horizontal and/or vertical visibility may prohibit the use of pilotage or celestial navigation. Further, whenever unknown or inaccurately prognosticated winds are present, the extrapolation procedure employed in dead reckoning is subject to serious and sometimes fatal error.

Radiolocation methods generally are unique in that they operate, for the most part, independently of the weather. Sonic systems are also independent of weather, but their use is confined to very small distances of the order of miles. In a relatively few instances, however, the utilization of radiolocation methods is curtailed by the elements. These cases are caused by meteorological situations which favor precipitation static or corona discharge because of snow or dust particles; mechanical failure of radio equipment by violent buffeting of the aircraft in severe turbulence or pounding of the vessel in heavy seas; lack of signals through destruction of ground stations by hurricanes, typhoons, simooms, etc. Other cases not strictly terrestrial in origin are radio fadeouts caused by ionospheric disturbances, which in turn are brought about by sunspot activity. However, the proportion of the time during which radiolocation procedures are nullified by vagaries of the weather is but a tiny fraction of that when other navigational methods are rendered inoperative for the same reason.

Before leaving the field of general navigation, it would be well to emphasize one requirement basic to all positional or navigational methods. This basic requirement stems from geometry: one observation establishes a given constant coordinate which the navigator terms a line of position. A second observation defines a second line of position. Obviously, the intersection of two lines of positions fixes the navigator's position precisely. In some instances each line of position may be a great circle; however, no ambiguity is caused by the second nadir intersection since the navigator in general knows the area in which he is located. Position may be specified by latitude and longitude (x, y), by distance and bearing (r, θ), or, in the case of dead reckoning, by velocity and time (v, t).

Under the tremendous impetus given electronics by World War II, several original electronic navigational systems have been evolved during the past few years. Some of the systems thus developed utilized radar techniques such as pulsing and high-precision timing, and others employed only simple continuous wave transmissions. Although it is not the intent to describe fully all radiolocation systems, it will undoubtedly prove illuminating to note briefly the distinction between the major radiolocation classes. A radiolocation system may be classified generally as one of three fundamental types: range, azimuthal, or hyperbolic. A fourth type, composite, consists of any twofold combination of the previous three (Figs. 1, 2, 3).

The range type of electronic navigational system is one of the simplest types proposed. It relies on the same fundamental principle as that underlying radar: if the velocity of electromagnetic propagation is constant, the distance or range between two points is proportional to the time required for a radio wave to make a round trip between these points. Even if the coordinates of either the probing station or the remote point were known, this

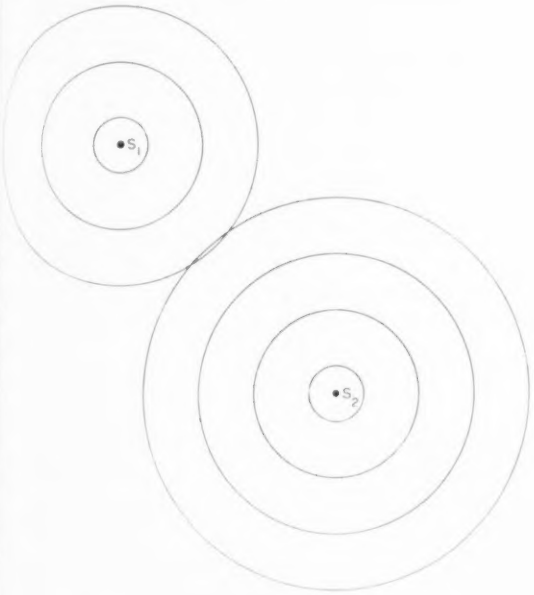


Fig. 1. Concentric circles about each range station. S_1 and S_2 indicate isopleths of constant distance between the observer and the range station. (Not all isopleths have been drawn.)

procedure establishes only the distance between them. In general, the navigator then knows that he is located somewhere on the surface of a sphere whose radius, measurably changing in a high-speed aircraft, is known. The intersection of this sphere with the surface of the earth provides somewhat

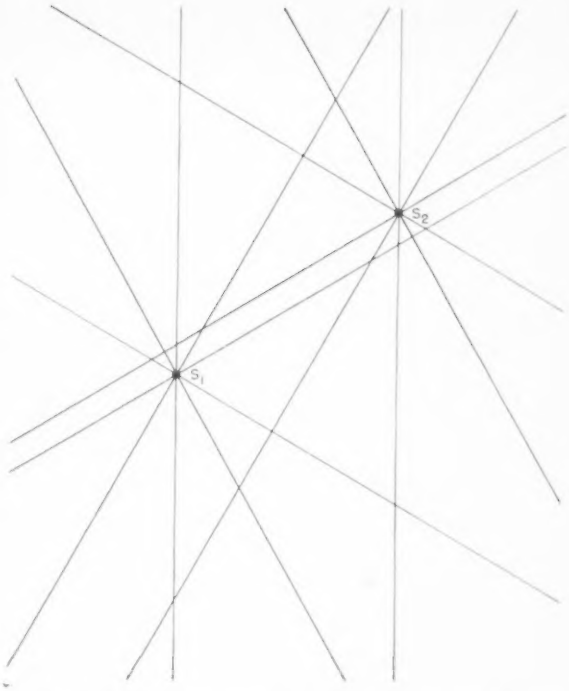


Fig. 2. Radial lines emanating from the azimuthal type navigational transmitters indicate isopleths of constant bearing (on a plane earth). Although the stations S_1 and S_2 are shown as a point, each may include several antennas usually separated by a small number of wave lengths.

concentric lines of position, but when these are drawn upon a map, their actual shape depends upon the map projection employed. The surface navigator, however, by maintaining a constant time delay on his receiving instrument could navigate along any desired "circular" line of position.

In some systems, such as Shoran and Oboe, the mobile craft emits the pulse and receives its reflection. An improved method would allow some means for distinguishing, from among the multitude of reflections produced by prominent objects in the path of the radiation, one particular echo from a reflecting point whose coordinates are known. This method, employing interrogators, solves the problem by installing responder-beacons, or Racons, at desired specific ground locations. Usually the range type of navigational system is restricted to a relatively small area on the surface of the earth by the limited power of the mobile transmitter.

Azimuthal navigational systems are among the oldest radiolocation aids known and are quite familiar through the widespread use of one of the methods, the typical direction-finding system. DF systems are based on the fact that the electromagnetic energy received in certain classes of antennas, square or circular loop antennas for ex-

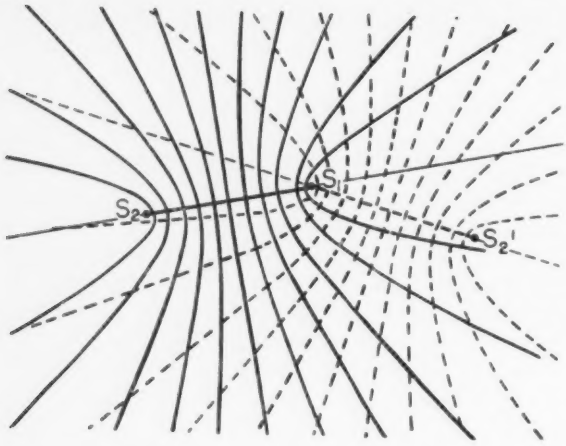


Fig. 3. "Hyperbolas" indicate isopleths of constant time difference between the time of arrival of signals from station pairs S_1 and S_2 and S_1 and S_2'.

ample, is a function of the mutual orientation of the radio wave front and the antenna. By rotating the antenna, the received signal strength varies correspondingly and thus provides a means of determining the angle between the incoming wave from the distant transmitter and the loop antenna. When a radio compass is available the preceding information may be manipulated to give the angle between the incoming wave and geographical north (after corrections have been applied).

Navigation by pure DF methods, however, may be deceptious, particularly at great distances from the transmitter. The deceptiveness arises from two facts: the meridians are not parallel but convergent, especially at high latitudes; and the great circle path—followed by the radiated signal from the transmitter—is not coincident with the line of position for a given DF bearing. At distances within about 150 nautical miles the great circle path and the line of position for a given bearing are, for practical purposes, closely identical (Fig. 4). For utmost accuracy, corrections must be applied for the convergence of the meridians and for the difference between the magnetic variation at the craft and the radio station. In following a given DF bearing, a navigator will describe a loxodromic curve rather than a great circle path.

Under the generic term of azimuthal systems, there may be classed a subheading called directional azimuthal systems. These more recent navigational networks require an electrical phase or amplitude measurement, resolved by visual or aural or possibly even metering methods, which is essentially converted into a bearing, or radial line of position. When the transmitter is a fixed ground station, the lines of position in the ideal case appear as great circles crossing at the station. With long wave lengths, the directional effects are obtained by the use of two or more antennas, which from great distances appear as a point, separated from a quarter to several wave lengths. At these low frequencies the resulting radiation patterns indicating relative signal strength may appear as lemniscates, cardioids, or 3- or 4-leaved roses, depending upon the antenna spacing.

In azimuthal systems, directional patterns at low frequencies are not produced by physically moving or rotating large cumbersome antennas, but by electronically switching to different pairs of fixed antennas, or by changing their electrical excitation. The transmitter usually emits at least two types of signals, which correspond to two different azimuthal directional patterns. Thus, regardless of the absolute intensity of the two signals at a given time and place, their relative strength is constant and is dependent solely upon the bearing of the observer with respect to the transmitting station.

Perhaps the most familiar example of an azimuthal system of this type is the common A (·–) N (–·) network (Fig. 5) covering the airways of the United States, Canada, and other countries. This system is received aurally by means of a radio receiver. The installation suffers the disadvantage of being limited to four directions. (Under special circumstances sporadic, unplotted, or undesired courses arise which may result in disaster.)

Another example, which, however, is not restricted to only four directions, is the Consol system. This and similar directional azimuthal systems do not provide complete 360° horizontal coverage about the antenna array, inasmuch as the limited sectors (usually in the vicinity of the base

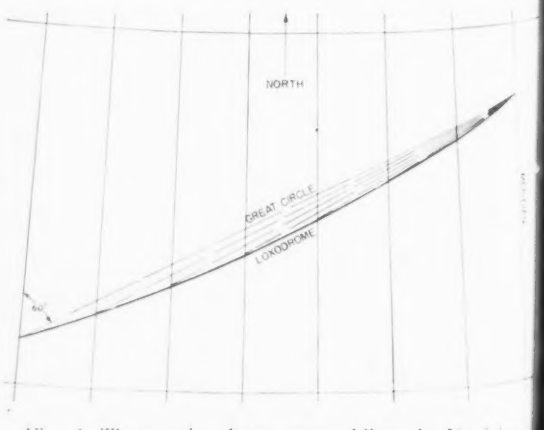


Fig. 4. The spacing between meridians is 2° of longitude. Even at the comparatively small distance portrayed above, appreciable deviations between the DF bearing and the rhumb line are evident. At greater distances from the DF station, deviations increase rapidly.

line extension) the error in position determination becomes excessive. At the transmitter a pattern of dots and dashes is radiated through the combined means of special antenna arrangement, sequential transmission of distinct radiation patterns, and uniform phase shifts of the antenna current. To the navigator receiving them, these signals appear as a characteristic pattern of dots, dashes, and equisignal zone (hum), the pattern being a function of the bearing of the navigator with respect to the transmitter. Using as an argument the number of dots and dashes received, the navigator refers to special charts and quickly determines a line of position.

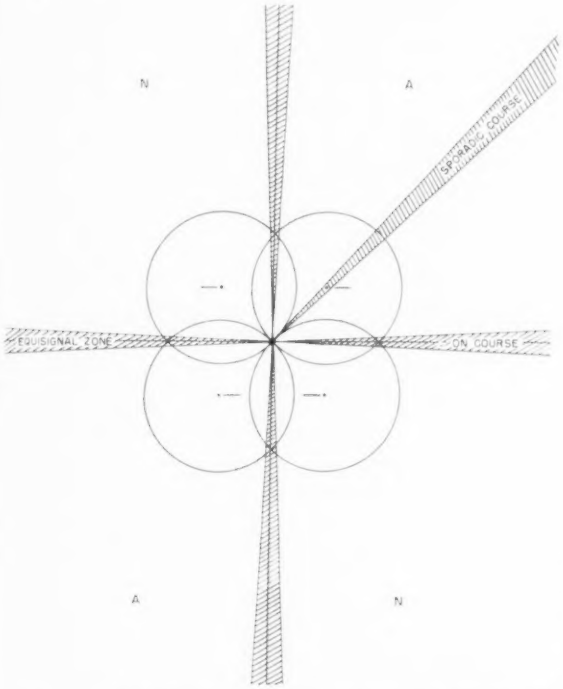


Fig. 5. The four perpendicular courses are typical of those found on *A–N* airways ranges. Sporadic courses may arise because of mountains, peculiarities of the weather, or discontinuities in ground conductivity.

Still other systems provide omnidirectional (360°) coverage about the transmitting station. Two such types are Navaglobe and CAA Omnirange. In the former the bearing indicated at the mobile craft is the bearing of the observer from the ground transmitting station.

In a hyperbolic navigational system the difference in time of arrival of two radio signals is measured. The definition of a line of position is similar to the definition of a plane or spherical hyperbola. A plane hyperbola is the locus of a point which moves in such a fashion that the absolute value of its distance from two fixed points equals a positive

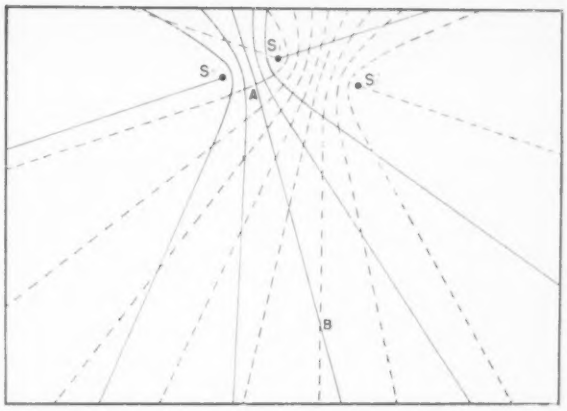


Fig. 6. Absolute accuracy in position determination is greater near point *A*, where the isopleths intersect almost perpendicularly, than near *B*, where the isopleths intersect at a relatively small (acute) angle.

constant whose value is always less than the distance between the foci. If the velocity of electromagnetic propagation is assumed constant, a hyperbolic line of position in a plane surface would be defined as an isopleth of constant difference in arrival time of two synchronized signals which had been emitted one at each focus, respectively. In actuality, since the earth is a spheroid and since the radio waves propagate in three dimensions, a line of position in a hyperbolic navigational system lies on the intersection of a hyperbaloid of two nappes with the geoid.

It is thus evident that geometry and the measurement of time difference are the qualifying factors for a hyperbolic navigational system. Either a pulse or phase technique may be employed. With a pulse system, such as Gee or Loran, the difference in arrival time of two pulses provides a line of position directly. In the phase method (Decca or Raydist), a comparison of the relative phase of the two continuous synchronized wave signals emitted, one at each focus, must be obtained. One great advantage of hyperbolic and azimuthal over range system types is that in the former methods no signal need be radiated from the aircraft. A simplified hyperbolic system is portrayed in Figure 6.

In a composite electronic navigational system, each of the two stations is of a different type. For example, one station may be of the range, the second of the azimuthal, type. The resulting pattern of lines of position may be seen by superposing the appropriate diagrams (Figs. 1, 2, 3). Thus far, however, only the simplest type, range and azimuthal, has been inaugurated.

It is appropriate at this time to mention that an inherent accuracy factor existing in each of the

various radiolocation or other types of navigational systems is predetermined by the angle at which two lines of position intersect. Obviously, any system which could provide an orthogonal relationship between the lines of position determined by the navigator would contain the greatest accuracy factor in this respect. Conversely, if the system provided lines of position which crossed at small angles, the accuracy would be materially decreased (Fig. 7). Orthogonal radiolocation systems might be devised by combining range and azimuthal types of stations which are mutually perpendicular. The hyperbolic type would require a family of orthogonal "ellipses"—a requirement difficult of fulfillment. As it has a comparatively small service area, a range type of station has limited usage in a long-range navigational system. Nevertheless, without such stations in a long-range navigational system, the principle, or orthogonality, in the intersection of lines of position is sacrificed.

A brief summary of the various systems and some typical examples of each type are given in Table 2. (The list is not intended to be exhaustive.)

TABLE 2

ELECTRONIC NAVIGATIONAL SYSTEMS

Range type	Oboe
	Shoran
	Racon
Azimuthal type	*A–N*
	Direction finding
	Sonne
	Consol
	Popi
	Navaglobe
	Omnirange
Hyperbolic type	Decca
	Raydist
	Gee
	Loran
Composite type	Range and azimuthal (1 station of each type)
	Range and "hyperbolic" (1 station of each type)
	Azimuthal and "hyperbolic" (1 station of each type)

Various frequencies could, of course, be employed for any of the types listed.

From the partial register given in Table 2, it is obvious that a fair choice is allowed in the selection of station types and systems needed to fulfill any particular navigational requirement. For each station type employed, however, the navigator must have aboard his craft, and must know how to operate, the proper receiving equipment. (Some systems, such as *A–N* or Consol, require little other than a receiver and earphones; others, such as Loran, require highly specialized apparatus and skill in their use.) In addition, each type of system is usually accompanied by its own series of maps printed with lines of position pertaining to the system. If various adjacent countries installed and operated diverse navigational systems,

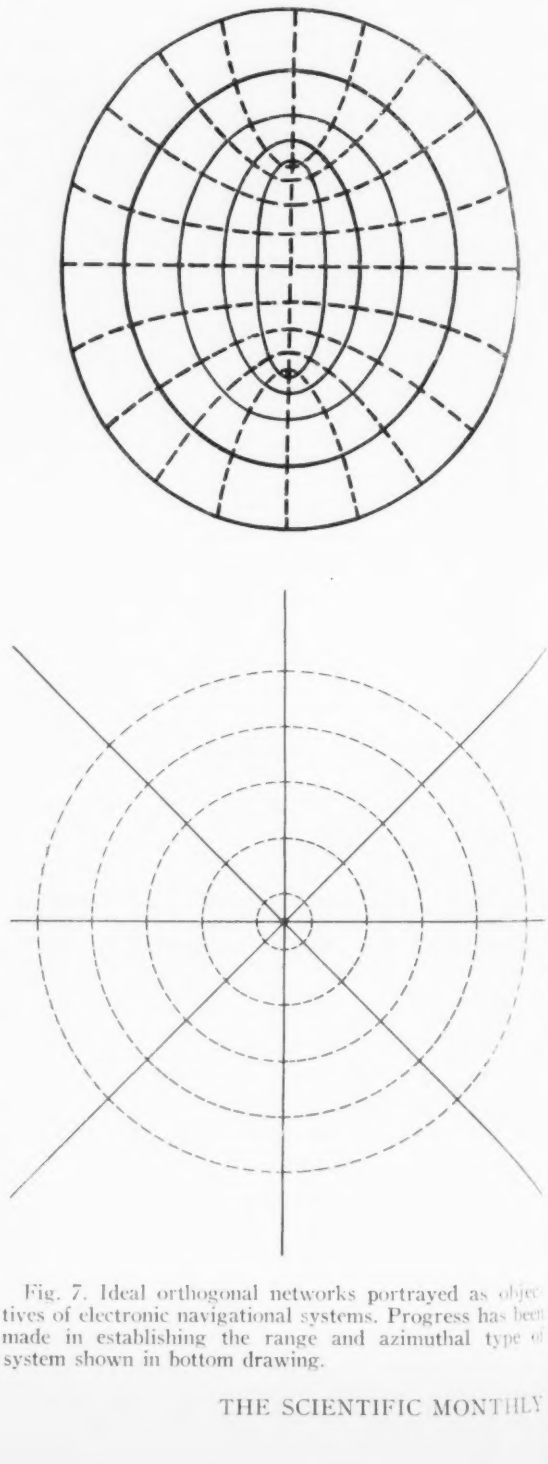


Fig. 7. Ideal orthogonal networks portrayed as objectives of electronic navigational systems. Progress has been made in establishing the range and azimuthal type of system shown in bottom drawing.

long range vessels or aircraft would be required to invest in excessive installations of equipment, to keep on hand large numbers of charts, maps, and books of instruction, and to train their navigators in the use of a wide variety of equipment, all to accomplish a similar purpose. Obviously, a standardization of navigational systems was necessary to avoid confusion, large investments, and displacement of pay loads.

To effect such a standardization, delegates from nearly two thirds of the member countries of the then Provisional International Civil Aviation Organization (PICAO), the society now known as the International Civil Aviation Organization, (ICAO), met in Montreal during October and November 1946. In addition to other recommendations, the PICAO Conference recognized the value of standard- or medium-frequency Loran as a long-range navigational system. Standard Loran already was providing adequate coverage over several important oceanic regions, and it was deemed of sufficient worth to warrant further installations as the demands arose. However, notwithstanding its many favorable aspects, standard Loran had two serious deficiencies: it would not provide a coverage area over land because of the large signal attenuation at its frequency, and it would not provide complete transoceanic coverage within its accurate but somewhat limited ground-wave service area. The same system operated on a frequency about one tenth that employed in standard Loran seemed to overcome both objections. Low-frequency Loran thus appeared to offer the greatest promise of meeting the desiderata of a long-distance navigational system. It also met the problem of instrumentation by not requiring completely new equipment, but merely a small frequency converter, which allowed LF Loran signals to be received on standard Loran equipment. Thus, with the converter, both standard and LF Loran signals could be utilized with, in many cases, already existing apparatus. The conferees were sufficiently impressed with the range and accuracy characteristics of LF Loran to recommend tentative installation dates (Table 3). This step is the first one toward a global standardization of long-range electronic navigational systems.

In view of the importance of standard and LF Loran as geonavigational methods of the immediate future, these systems will be described in some detail. Operationally, both are very similar except for the additional advantages of LF Loran caused by the enhanced propagation characteristics at the lower frequencies.

"Loran" is a word coined from the initial letters of its definitive phrase: LOng RAnge Navigation. The basic proposal for Loran was devised in 1940, and in 1942 the first fixed stations were placed in operation along various coastal regions. (Simultaneously with the development of Loran, the Telecommunications Research Establishment, of England, developed a similar system, Gee. In 1942 information between Radiation Laboratory and Telecommunications Research Establishment was exchanged, and research on both systems continued. Gee was primarily intended to be employed by high-altitude aircraft for a comparatively short range.)

The operational frequency of standard Loran lies in the vicinity of 2,000 kc/s, and it may be employed by both aircraft and ocean-going vessels. At these frequencies, radio waves propagate around the curvature of the earth for fairly large distances. Although the general principles of "hyperbolic" navigation have already been presented, a Loran system in more detail is shown in Figure 8. At some point P a navigator measures the difference between the arrival time of two pulses, one radiated from the system master station S_1 and the second radiated from the system slave station S_2. A similar time difference is established for the second system pairs S_1' and S_2'. Each time difference corresponds to a line of position for the particular station pair involved. In practice, master stations for adjacent Loran transmitter pairs are

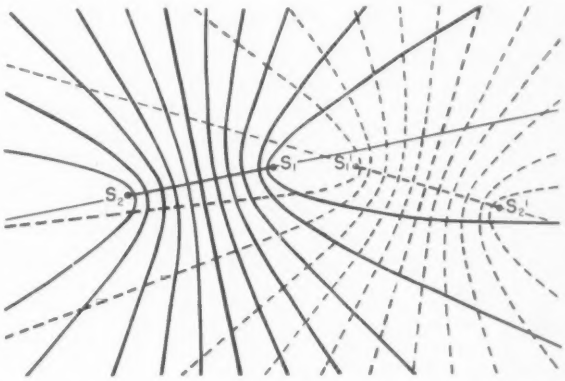


Fig. 8. The master station transmitters are indicated by S_1 and S_1', and their associated slave transmitters by S_2 and S_2', respectively. (Simplified plane Loran system.)

TABLE 3

Region	Tentative Installation of LF Loran
North Atlantic Ocean	January 1949
South Atlantic Ocean	January 1951
Southeast Asia	January 1951
African continent	January 1951
Australasia	January 1951
Pacific Ocean	January 1951

usually coincident; thus, S_1 and S_1' would be combined and the station termed a double master.

Instead of spending considerable time in making stellar or solar observations, and performing intricate calculations, the navigator now makes two simple measurements of time difference. By noting the intersection of two isopleths bearing these values on a map, the navigator locates himself. Some means of distinguishing the signals from the above two station pairs must be given the navigator. The identification may be easily accomplished by allowing one pair of stations to transmit pulses at a rate of perhaps 25 pulses per second, whereas the other pair transmits at a slightly different rate, such as 25 and a fraction pulses per second. A means is also provided to warn the navigator when difficulties at either transmitting station prevent the accurate transmission of pulses.

If pulses were emitted by each station simultaneously, the system would contain an ambiguity because of the symmetrical spacing of equal-valued lines of position on either side of the baseline perpendicular bisector. To remove this uncertainty, the slave station radiates its radio pulse, identical in wave form with the master pulse, at a constant delay period after receiving the master station pulse (Fig. 9). The constant delay is equal to 22,000 microseconds, plus the time of transit of a radio wave between the two stations in microseconds. By an electronic arrangement in the receiving equipment, 20,000 microseconds are automatically subtracted from the navigator's observation. The remaining 2,000 microseconds and the transit time between the two stations are absorbed in the method of labeling the isopleths of constant time difference on the Loran navigational chart. This chart has, in addition to ordinary latitudes and longitude lines, topographic features, lakes, rivers, etc., Loran lines of position printed in different colors for every station pair.

At the receiver, the pulses from each station are rectified and displayed upon the screen of a cathode-ray oscilloscope. By introducing a time difference equal in magnitude to that existing between the pulses when they impinge upon the antenna, both master and slave pulses may be superposed. The introduced time difference may be precisely determined by radar timing techniques incorporated in the equipment.

The accuracy inherent in any navigational system is a major consideration in its ultimate adoption. In Loran, the accuracy is not constant, but varies over the service area. In the most favorable regions, along the base line where the lines of positions converge, an accuracy to within a few

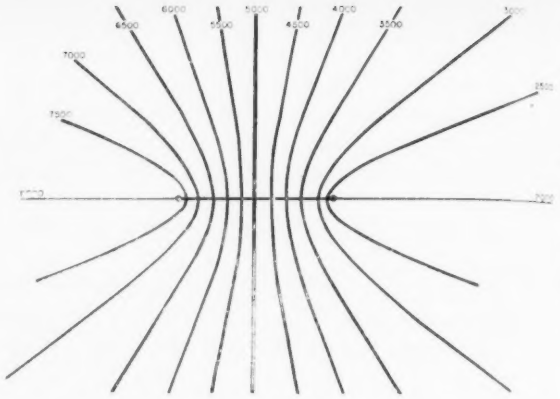


Fig. 9. By inserting a residual coding delay of 2,000 microseconds in the transmissions from the slave station, the ambiguity in position determination that would otherwise exist is entirely eliminated. (Plane Loran.)

tenths of a nautical mile or less is possible. As distance perpendicular to the base line is increased, the lines of position become increasingly divergent, thus lessening accuracy. Also, in the region of the "degenerate hyperbola," i.e., along the baseline extensions, accuracy is so poor, because of the wide separation between lines of position, that normally this region is not used for position determination.

A second factor affecting system accuracy is the distance between the master and slave transmitting stations. If the base line is doubled, the error of position determination at any particular point under identical conditions is approximately halved. Still a third factor is the principle of orthogonality, which with Loran chains cannot be maintained. By a judicious choice of station location, it is possible to have the lines of position from two adjacent Loran systems closely perpendicular in some small, particular area. In general, however, especially at outlying distances from the stations, the angle of intersection of the lines of position will be small. Thus it is evident that the accuracy within a Loran service area is intimately associated with the navigator's position with respect to both transmitters. A rough estimate of accuracy is contained in the general empirical rule that the accuracy of position determination is about 1 percent of the mean distance from the transmitters. If the navigator is 1,000 nautical miles from the stations, he is located within about 10 nautical miles of his determined location.

The coverage of standard Loran over sea is about 600 nautical miles during daylight and about 500 nautical miles (but variable) during darkness. The lower coverage area at night is principally caused by the increased noise, which is of a greater relative intensity because of the enhanced propa-

gation of electromagnetic waves. The decreased signal-to-noise ratio correspondingly decreases or eliminates the possibility of obtaining an accurate Loran observation. The smaller coverage area at night may be almost quadrupled by employing pulses which have been refracted from the ionosphere. In this procedure, a sky wave is employed which leaves the earth at a low angle at the transmitter site, strikes the E layer, and is refracted back to earth. This pulse is known as the first-hop E. By employing first-hop E pulses the coverage area during night is increased to about 1,300 nautical miles from the transmitter. The additional transit time required by the signal is taken into consideration by a series of correction tables. Generally, little confusion exists in distinguishing between first-hop E and ground-wave pulses. Since the height of the E layer is not constant, the resultant accuracy is decreased somewhat. The final results are still satisfactory, however, and during the war some blind bombing was effected by using sky waves. When master and slave stations employ sky waves for purposes of synchronizing their transmissions, the system is known as Sky-Wave Synchronized, or SS Loran.

The simplicity and many advantages of standard Loran led to its installation over large coastal regions of the globe. Its use, however, is distinctly restricted to oceanic or other water areas. It does not provide satisfactory navigational service over land masses. In this respect it is inferior to the Gee system. Another disadvantage, or perhaps disappointment, in standard Loran is its inability to provide a range of at least 1,000 nautical miles under all conditions.

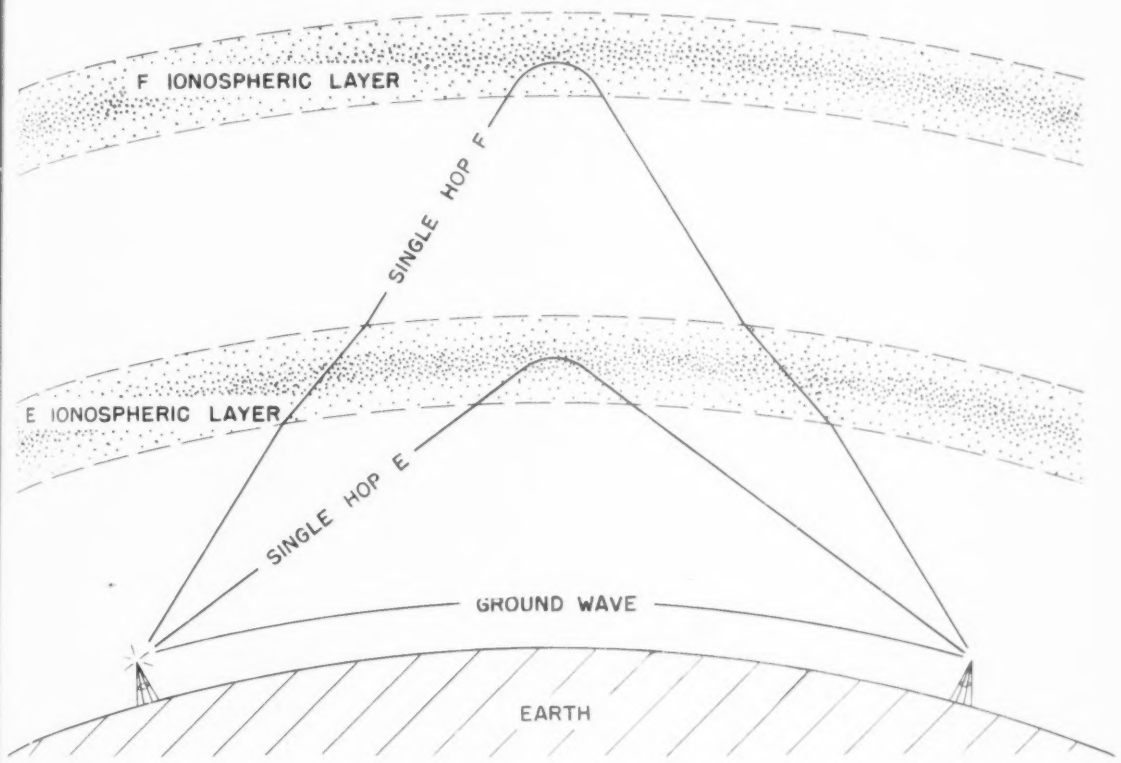


Fig. 10. Because the ground wave attenuates much more rapidly than the sky wave, the latter may still be employed at distances where the former is of such small intensity that it cannot be utilized. Either E- or F-layer reflections may be obtained, depending upon the radio frequencies transmitted.

The greatest factor preventing increased coverage for long distances over both land and sea was the carrier frequency being employed. Since absorption and reflection phenomena, ground-wave range, etc. are frequency-dependent, frequency considerations were re-examined and a lower one recommended. When a carrier frequency about one tenth that utilized in standard Loran was used, both overland coverage and increased ground-wave range were obtained. In addition, several other advantages were apparent. Because of the increased

range, the ground transmitters could be more widely separated. Thus increased accuracy over the service area was obtained. Also a not-inconsiderable monetary saving resulted because of the fewer number of ground stations required to cover a given area. Because of instrumental difficulties, accuracies with LF Loran are not as good as those obtained in standard Loran; nevertheless, they are comparable. One result of the theoretical investigations was that if four LF Loran transmitters were located at the corners of a square 1,000 nautical miles on a side, a root mean square distance error of 5 nautical miles or less was obtainable over a service area of close to 3 and 1.7 million square nautical miles by day and night, respectively. Basically, the operational principles and techniques of LF and standard Loran are similar. With LF Loran many difficulties of technical character made their appearance, however. One such minor difficulty may be mentioned—a quarter wave-length antenna is about 1,300 feet in height. At the present time this antenna is a thin copper-clad wire supported by a barrage balloon.

It should be emphasized that Loran and other radiolocation methods are confined to terrestrial navigation. With the advent of rockets and, in the near future, interplanetary flights, these systems would be of negligible value in comparison to celestial navigation methods. As long as speeds of less than 150 nautical miles per second occur, relativistic effects in time of transmission may be neglected. One disadvantage of all Loran systems is their unsuitability as homing systems unless a line of position passes through both the navigator's position and his objective.

TIME

If time were but a constant pendulum
Swinging between the boundaries of day
And night, its iron regularity
Would drive fear's midnight messengers away.

But time's a seesaw. With a masked unknown
For fellow-passenger, who can foresee
Whether he'll tip the balance toward the sun
Or the cold earth's rock-strewn topography?

FRANCIS BARRY

THE EVOLUTION OF TERRESTRIAL LIFE AS SEEN BY A BIOCHEMIST

ADDISON GULICK

Born in Osaka, Japan, of missionary parents, Dr. Gulick received his Ph.D. in 1910 at Würzburg. He taught briefly at the Universities of Toronto and Minnesota; since 1912 he has had charge of physiological chemistry at the medical school of the University of Missouri, where he is at present chairman of the Department of Biochemistry.

TO THE naïve, terrestrial life may seem the least involved and the most natural way in which to live, but from the physiological and biochemical viewpoint, such life is a highly specialized evolutionary adjustment to a difficult situation. Historically, the difficulties have proved to be so severe that sometimes tens of millions of years have been involved in the process of gradually overcoming them.

It is universally agreed that life was originally aquatic, and that the first great expansion of life forms was in the sea or possibly in shallow waters bordering the sea. Nearly all the mineral nutrients of living organisms are extremely water-soluble. That places them among the substances that leach out from the eroding rocks and are thus carried in a dissolved form by the rivers into the ocean. We do not know how long this process has been going on. It must have begun just as soon as the earth stabilized at a temperature that permitted rainfall. Geologically, the evidence would be the formation of sedimentary rocks. The date of the very earliest sedimentary rocks has not been determined and may never be. In one place and another among the very ancient sediments there are scattered conditions that permit an approximate dating of some particular formation. The most ancient sediments that have been determined by way of the so-called radioactive clock technique have been found to date back about 1,750 million years. That does not reveal what older sediments might exist or may once have existed and the records perished.

It was undoubtedly a long time from the first formation of rain water and seas upon the earth before the exceedingly difficult combination of coincidences that brought forth life may finally have occurred. Nearly all the very ancient sediments are by now so heavily altered that there is little in them to show whether life existed at the time of their deposition. The likeliest manner in which such evidence may survive would be through the existence of streaks of graphite caused by the accumulation of some sort of vegetable material. Such evidence can be traced with a fair degree of assurance some two and a half times as far back as the earliest classifiable fossils; that is, to about 1,200 or 1,400 million years ago, when the graphite-rich Grenville sediments of pre-Laurentian age started forming. But that fails to tell us what primitive, slowly evolving forms may have existed at earlier dates. Such deposits of carbon are usually a result of concentrations of plant life in shallow water.

By the time the leaching of the continental soils had been going on for hundreds of millions of years, it would seem obvious that these surface soils, despite many submergences and re-elevations, must have become very seriously depleted. We must further suppose that the river waters originating upon these land masses must also have ceased to contain much soluble material, so that there was already a sharp contrast between the nutrient-poor lands and fresh waters, on the one hand, and, on the other, the perhaps already overstrong brine of the ocean. Under these circumstances probably the earliest parting of the roads into unlike habitats was the division into marine and fresh-water organisms.

We cannot say positively, but only with great probability, that the first colonization of land came from fresh water. This would be the less difficult of the two alternatives because the mineral content of ground water and of fresh surface water are quite similar. Plants must necessarily have preceded animals as colonists on the land, because otherwise animals, completely dependent upon organic foodstuffs, could not have nourished themselves.

Vascular plants. It is not unlikely that the earliest organisms to establish themselves on land may have been microscopic green algae. This has to remain a surmise because plants of this sort have very little power to leave geological evidences of their existence. Presumably following them, but still at a relatively early date, would

have been the mosses, an essentially fresh-water stock which became the starting point for the development of types having strands of vascular or woody tissue, thereby forming the first links of the evolutionary chain that consists of higher mosses, ferns, and flowering plants. The earliest good fossils of vascular plants known thus far are credited to the Silurian of Australia—age about 330 million years—but the stock they represent must at that time have been long acclimated to dry land.

This great block of plant life, beginning with mosses and culminating in flowering plants, has considerable solidarity of type and is subject on land to a rather characteristic group of special limitations and potentialities. Their answer to the problem of water shortage was to run rootlets down into the earth. This method of nourishment establishes an essentially one-way flow of water. From the rootlets in contact with ground moisture to the green structures exposed to sunlight and to evaporation, living on a soil with low sodium and low chlorine content, their whole physiology shows adjustment to shortage of these more soluble elements. In fact, most of them have by now evolved to where more than a trace of sodium chloride amounts to a mineral poison. Regulation, through one means or another, of water and mineral intake was necessary, first, because many of the essential elements have been leached out until they are scarce, and, second, because if any form of mineral becomes absorbed copiously the excess can never be disposed of, and so can hardly fail to leave troublesome deposits in the leaf regions, where they are left behind when the water evaporates.

A rooted plant is unable to migrate, so for survival some form of dry dissemination is necessary. Primarily, we observe in mosses and ferns a system of dispersal by dry spores. Much later comes the dissemination of seeds by animal agency or wind, and of pollen by visiting insects.

When dissemination has come to be a dry process it has to be dissociated from fertilization, which in its earlier evolutionary forms is an aquatic process. Mosses and the principal fern forms are still aquatic with reference to reproduction, to the extent that they require at least a drop of water in which the spermatozoid can find the egg cell. It is only in flowering plants that the mechanism known as a pollen tube completely supersedes the original aquatic pattern of their life cycle.

Another adaptation in plants meets the difficulty of dry air and the resulting high rate of evaporation. From the liverwort mosses upward, nearly all plants provide themselves with minute air spaces, enclosed within their tissues, whereby the actual individual living cell is removed from contact with external air and allowed to function in the environment of a water-saturated atmosphere.

Excretion also is a special problem coming from terrestrial life. Aquatic forms, particularly when simply organized, can get along with relatively primitive methods of disposal of the excretion products. On land the nonvolatile excretion materials must usually be especially provided for. The green plants are less troubled by this problem than are the other land organisms, because they are well supplied through sunshine with the wherewithal for reactions that absorb energy. This gives them the power to re-employ almost any form of waste material in biological syntheses. Plants have a notable ability to deposit by-products of their activity in the vacant cellular cavities or physiologically dead portions of the plant structure. It is hard to tell just what the explanation of some of these deposits may be, whether they are excreted end products from biological processes that have hit some chemical blind alley, or whether such deposits may actually be fulfilling some highly valuable purpose. The pitch in the nonliving portions of a pitch pine tree, for example, may conceivably represent such a "blind alley," but at least equally well it may be a provision that secures against the rotting of the tree trunk. Perhaps the most difficult material of excretory nature for the plants to dispose of would be the mineral material that arrives absorbed along with the necessary water intake. Apparently all adjusted land plants have learned to hold such unuseful mineral intake down to a very low minimum.

One might look upon the higher plants as representing the most perfect and most complete case of biological adjustment to nonaquatic life. One of the evidences in favor of this statement is that no other group of land organisms begins to compare with such plants for ability to live in the absence of sodium chloride, the salt which more than all other salts has been effectively removed from the land areas of the world.

Insects and arachnids. Although it is impossible today to decide what group has the distinction of being the earliest animal type to have made its way out of water onto the land—whether the arachnids, comprised of scorpions, spiders, and their relatives, or the centipede tribe, or the insects, or the latter two groups together by way of a common early land ancestor—the best-docu-

mented proof of antiquity among land animals goes to the arachnids. This is because the earliest of them were scorpions, a form that is relatively well adapted to leaving fossil remains. Spiders were a later development, arising by way of a slender-tailed transition form. Scorpions are found on land in the Silurian strata, approximately 330 million years ago. They show obvious relationship to their contemporaries and predecessors, the marine Eurypterids, which were an ancient stock related to the crustaceans, and ancestral to the present Limulus, or horseshoe crab.

When they became terrestrial, the scorpions enclosed the eurypterid book gills behind spiracles, and so produced "book lungs," from which in due course the spiders developed a tracheal system, superficially similar to those of insects. But it is unlikely that this resemblance strikes any deeper, as there is no indication that insect ancestors ever had book gills.

The origin of insects is much more obscure. The winged forms, which can be traced back 250 million years into the time of the coal measures, were clearly descended from wingless terrestrial predecessors, which are thus far first discoverable in the Devonian, aged about 300 million years. The ancient wingless insects were very small, almost microscopic in fact, and of extremely lowly organization. Fortunately, their rather near counterparts are still to be observed in the ground litter of present forests, and the comparative anatomy of their soft parts can be studied. In this way it is possible to establish a probable relationship between them and certain small, primitive kindred of the centipedes. Hence comes the hypothesis that insects and the centipede tribe had a common origin from a very primitive arthropod type that was somewhat collateral to the crustacea. They may even have emerged onto land before becoming bulky enough to require special respiratory organs in either the aquatic or the aerial environment. Respiration by spiracles and tracheae gives the impression of being a joint inheritance of insects and centipedes from a very early common terrestrial ancestor.

Insects, centipedes, and arachnids all have exoskeletons, inherited from their aquatic ancestors. They all have their physiology founded on similar solutions of the problem of air breathing, and of reproduction by eggs under dry conditions. Thus the different groups exemplify in several respects a remarkably parallel or convergent evolution.

Life within an exoskeleton has no great disadvantages for a crustacean that can float quietly in the water, uninfluenced by the distorting effect of gravity during the periodically repeated occasions when for a few hours after its molt it simply has no skeleton. On land it is a different story; an insect cannot afford to grow so large and heavy that its weight will ruin its shape while it is soft. As competition on land increases, natural selection will enforce the limitation on size more rigorously. Thus, all told, modern insects are on the average small organisms. Even in ancient times their maximum observed size was far below the corresponding maxima for vertebrates—a roachlike form with about a 2-foot wing spread and a dragonfly with a 29-inch spread being the largest.

The laws of physics that affect a creature as small as the average modern insect are quite different from those that rule our own daily lives. The muscular power to jump or pull is measurable by the cross-sectional area of the muscle involved —a two-dimensional quantity—whereas body weight and weights of load carried are three-dimensional. Because of this, an insect with linear dimensions 1/100 of a man may have, in these terms, a muscle strength 1/10,000 of the man's and a body weight 1/1,000,000 that of the man—roughly, a one hundredfold relative mechanical advantage over the man in the ratio of strength to weight. On the other hand, some physical forces that are completely negligible for us may have ominous consequences for the insect. Any small insect's very life is threatened if it has the misfortune to suffer a moistening contact with a good-sized dewdrop, since surface tension is a force that far outclasses his muscle power. Hence their need for various devices to combat moistening with water—special waxy substances on their surfaces, water-repellent hairs, wing-scales, etc.

As natural selection pushes organisms gradually to higher and higher levels of muscular activity, the oxygen needs of their tissues increase in similar proportions. In many modern insects the branching of the tracheal system which supplies them their oxygen has been carried to astonishing extremes, and the ultimate ramifications become tubules only 2–3 μ (thousandths mm) in diameter, which actually penetrate some of the cells that they serve for gas-exchange purposes. These fine-bored air tubes bring up a new and unexpected complication. Their walls are semipermeable membranes, and but for the aid of the osmotic pressure of the surrounding body fluid, the capillary force, which grows progressively stronger as the tubes grow narrower, would suck enough water into the tracheal tubes to waterlog them and completely destroy their function. The smaller the tubes, the

higher is the salt content required in the body fluids to prevent this waterlogging. Thus we find widespread among insects a salt content sufficient to give osmotic pressures 50–100 percent above that in mammalian blood. Usually potassium is the predominant kation; among anions, HPO_4 is more prevalent and Cl less prevalent than in the vertebrates. The comparatively smaller proportions of the ingredients of common salt might be considered an indication that during their long stay on land the insects have developed a more complete adaptation to the sodium chloride-impoverished terrestrial chemical environment than any vertebrate has achieved.

The laying of subaerial eggs, and the period of active biochemical metabolism while encased within the closed system of the insect pupa influence very powerfully the problems of the biochemistry of excretion. Aquatic animals, at least if not too large and complex, have only a relatively simple problem with their excretions, one that can be taken care of by providing opportunity for the outward diffusion of the useless end products of oxidative breakdown. On land, an adult animal may have to secrete water that will serve as the medium for the removal of wastes. The embryo inside an eggshell cannot even do this and, in order not to vitiate its available fluids, must find a relatively insoluble chemical form in which the excrement may be deposited. This is believed to explain why, among egg-laying land animals, the typical major form of nitrogenous waste product is almost universally uric acid, a substance involving a much more complex synthetic procedure than urea, but easily excreted as a semisolid porridge, or deposited harmlessly within the egg shell. As far as investigated, the insects illustrate this, as do also the spiders, reptiles, birds, and terrestrial mollusks.

Eggs with impervious shells and dry surroundings make it quite impossible for external insemination to occur, as was so easy in the water. Thus it is necessary for the first true land insects to have come from aquatic forms that had already developed the reflexes of copulation, along with a nervous system adequate to be the seat of such reflexes.

A major means of dispersal among marine forms is provided by tides and other water currents. The first colonists on land must have forefeited all such advantages. Later, this was retrieved again by insects acquiring wings, and by young spiders learning to float themselves on strands of web, let out to the wind. The history of the development of insect wings, the earliest of a series of cases of such evolution in several diverse animal types, is so ancient that its story is a lost chapter of paleontology, and the prospect of its recovery is not very strong. There have been various hypotheses, but little that can be said positively, beyond the fact that the opportunity for starting such a career must, as in all subsequent cases, have resulted from a habit of climbing tall plants or trees. We can no longer tell with certainty whether the primary advantage gained was escape from predators, or greater ability to function as a predator or forager, or to find a mate, or simply a better opportunity for the dispersal of the species. The fact that wings are only present after the last molt suggests the likelihood that for most insects their evolutionary significance is principally for the propagation and dispersal of the species.

Mollusks. The terrestrial mollusks all belong to the gastropods, or snails, a group whose marine history extends through almost the entire 500 million years' duration of the total fossil record. Since the fundamental classification of gastropods is based on the soft parts, and since bewildering similarities of shells often occur in diverse branches of the snail tribe, it is often difficult to appraise the fossil relationships. Forms believed to be land snails occur from the coal measures, 240 or 250 million years ago. Anatomical and physiological characters of their modern representatives make it appear that the derivation of the majority is from fresh-water pond snails. The latter, mostly relatively delicate and hard to preserve, are only known from much later strata, beginning with the Jurassic, less than 150 million years in the past. Cyclophoridae and Helicinidae, groups of land snails whose soft parts show affinity to entirely different branches of fresh- or brackish-water snails, must be credited to independent landward immigrations. The Cyclophoridae can be traced, apparently as land forms, as far as the Cretaceous, 100 million years or a little less, the Helicinidae at least to the earliest Cenozoic, about 60 or 70 million years.

No mollusks are able to sustain activity in a dry atmosphere, because aside from the shell they have no impervious outer covering. They can live only as long as they can maintain a moist skin. Except in rain forests and similar locations they are compelled to retreat into their shells or get under cover of ground litter or the like for the greater part of the daylight hours. At such times the Cyclophoridae and Helicinidae close themselves almost hermetically behind an operculum, such as their marine littoral cousins use. The Clausiliidae of

Europe and Asia retreat behind a close-fitting spring clap valve of shell substance, and the more typical Helices establish a fairly resistant membrane of dried mucin. Among the Endodontidae—more primitive kindred of the Helices—*Pyramidula alternata* impregnates its mucin barricade with a special bitter-tasting, brown-colored secretion that presumably serves as a repellent to marauders. Shell-less forms (garden slugs, etc.) go deep under the ground litter, doubtless depending on their greater facility at burrowing their way down; there, by means of their own slime, they protect themselves from desiccation.

From the limitations of their skin physiology, one might argue that snails have at best only halfway fitted themselves to life on land. Yet it has been found that, like insects and reptiles, the Helices have adapted themselves to the problem of terrestrial egg-laying by having a uric-acid pattern of nitrogen excretion. This should count as testifying to a long evolutionary process that has carried very far the essentials of adjustment to dry land. We should also note that very few animals can make out better in an arid climate than can some species of snails, because such a climate seldom fails to have dew at night.

Other invertebrates. It would be a futile task to try to discuss all the invertebrate types that have had greater or less success upon land. For many the fossil records are simply a hopeless void—such, for example, as the giant earthworms of the circum-Pacific area which, two feet long and half an inch thick, although entirely boneless, have the muscular strength to race through a canebrake like an agile snake. Just one more invertebrate group, the land crabs, seems worth mentioning, because it illustrates the relatively little-used marine-beachcomber road for entry to the land habitat. In various parts of the world the land crabs seem to have followed independently this opportunity for finding a new field for a livelihood. But in no instance have they become entirely independent of the sea water. From the lack of contrary reports, we assume that at least their larval stages still have to be completed in that medium.

Vertebrates. The vertebrates were somewhat tardy arrivals on land as compared to arachnids and insects. Just how they compared with other invertebrates would be more difficult to tell. The beginnings of the landward progress were through freshwater amphibian life, starting in the Devonian, perhaps 280 or more million years ago, becoming better developed in the Carboniferous period immediately following. If we traced the details of the changes involved, far too much of our report would be mere excerpts from comparative anatomy. But aside from that, there is still something to be said.

For vertebrates there were several physiological difficulties to the completion of the adjustment to dry conditions. Despite their fair size and somewhat elaborated nervous system, they may be supposed to have laid eggs that were outside-inseminated, and they doubtless depended upon aquatic life for the embryonic and larval stages. Their excretory system was presumptively based on urea, as is the case with modern amphibia, making it difficult to incubate eggs anywhere but in water. Modern amphibia have a soft moist skin, which limits their acclimation to land. The same may or may not have been true of the ancestral amphibia. Great dependence on sodium and chlorine, and a need for at least traces of iodine, may to some extent have constituted further difficulties, though the life in fresh water must have provided some training in getting along on a short supply of these substances. The difficulty that mammals experience with these shortages may come partly from their high turnover of materials and hence high losses, necessitated by a high-heat metabolism. It is possible, therefore, that the more sluggish ancestors of the reptiles did not find this problem so troublesome.

On the side of advantages we may list the possession of a swim bladder that can do duty as a lung; the well-developed vertebrate kidney, capable of evolving to carry out its new terrestrial duties, of taking over the whole burden of nitrogen excretion, and of developing control of the osmotic pressure and blood fluid acid-alkali balance in the face of more arduous conditions; the ability to produce keratin, or horny tissue, which the amphibian animals had not entirely lost, and which had prime value as an impervious external integment; and, last, the great mechanical advantage of an internal skeleton capable of growing as the body grows.

If reptilian hard parts can serve as evidence of reptilian physiology, we may infer that near, or shortly before, the start of Permian time, something like 200 million years ago, the difficulties were essentially overcome—dry skin, internal insemination, dry-shelled eggs, uric-acid excretion to conserve egg fluids and body fluids. The geological clock seems to indicate that it took 80 million years to progress from the start of amphibian life to the completed terrestrial form. A major aspect of the physiological change was the enhanced importance of the kidney. A frog or a

goldfish has all-over contact with its environment. A frog both breathes and excretes all over, and its lungs and kidneys serve only to make up deficiencies and to carry out those phases of excretion that need to be accurately adjusted, such as the body fluid's acid-base balance. Even a goldfish excretes urea, etc., principally through the gills, and uses its kidneys largely to excrete surplus water after capturing from it all possible sodium chloride. But a lizard or a land tortoise has a watertight sealed exterior, and of necessity what the frog would let diffuse out through the skin must now be placed under control of the kidney.

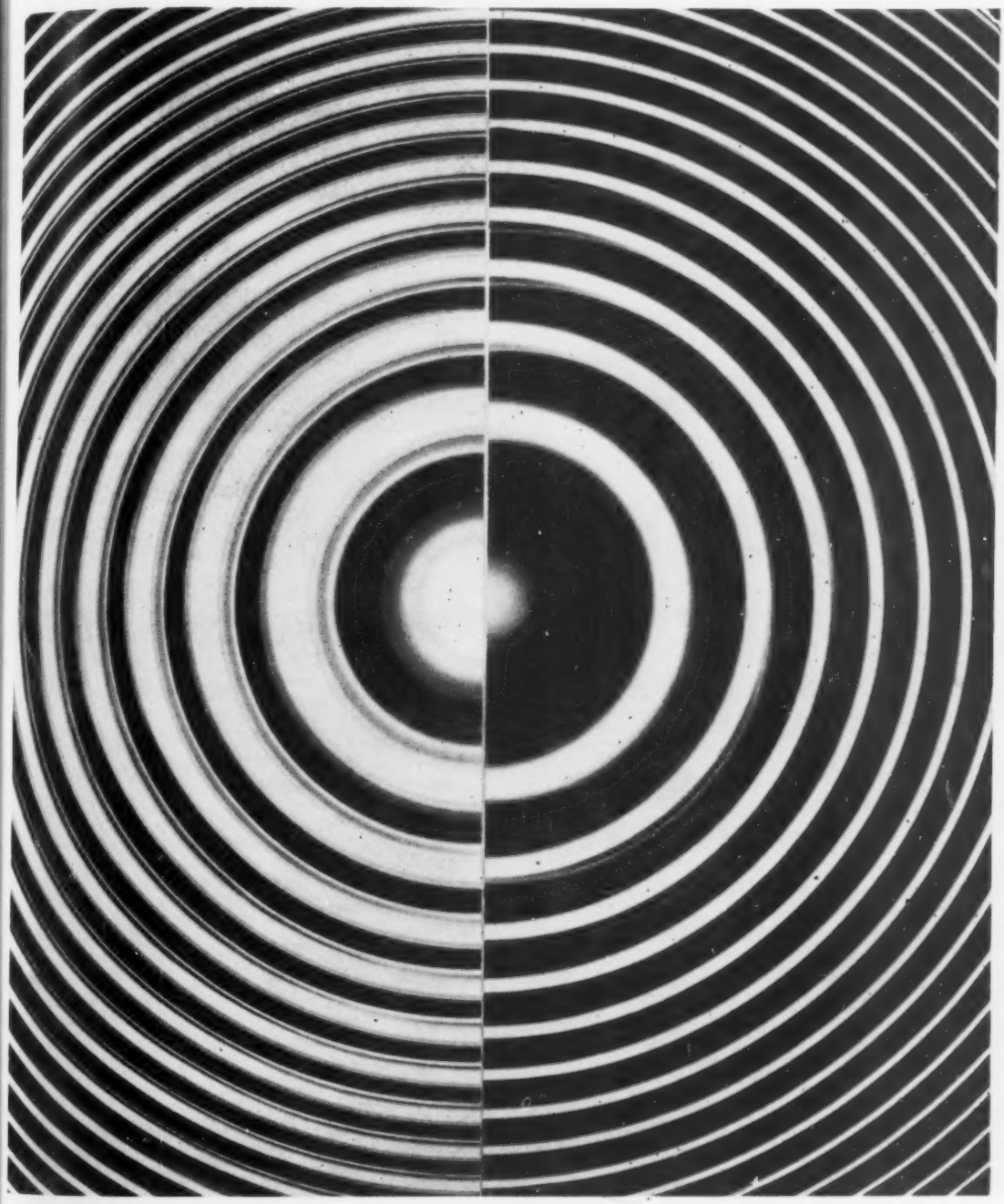
In a sense, the reptiles completed the physiological processes of adaptation to land, but once this was accomplished the doors were opened for a further set of related lines of advance. The exigencies of physiology within the enclosed egg led to creation of a peculiar set of membranes for the embryo, and this led next to the possibility of intra-uterine development by the altered use of these membranes. This was the line of advance followed by mammals, beginning with the multituberculates, related to marsupials, that came to light first in the Triassic, dating some 170 million years in the past.

When animals were finally completely freed from a watery environment, with its powerful control over the temperature of bodies immersed in it, and had become at home in a medium that supplied more than tenfold the quantity of oxygen gas per unit volume of the medium, the way was opened thereby for warm-blooded evolution. The first step in this direction had to be the correction of a deep-seated defect of circulation. The relation between vertebrate gills and circulation was such as to provide the body with fully arterial blood. Not so the pulmonary circulation in the amphibia and reptiles, since, being derived as a branch from the visceral circulation, it merely oxygenated a fraction of the blood and mixed it back with the rest of the stream. The evolutionary stages through which the heart and major blood vessels had to go to arrive at the four-chambered heart and the perfected circulation of mammals and birds is a fascinating story, but so familiar to every student of comparative anatomy that we cannot stop for it here. It is hard to tell how far to trust bone structure as revealing physiological organization, but, since in the Triassic the first mammalian skeletons are found and those of birds soon after in the Jurassic (140 or 150 million years ago), it seems easiest to place tentatively the arrival of warm-bloodedness at about the time interval of the Triassic to the Jurassic.

Next there follows, as a close corollary to the high level of metabolism and the high grade of sensory and neuromotor complexity, a renewed evolutionary emphasis on the nervous system—the trend that reaches its culmination in intelligent beings. Thus we see that the biological conquest of those relatively restricted fragments of the earth's surface that lie above water level was a major phase in the story of the evolution of living forms.

If a single cell, under appropriate conditions, becomes a man in the space of a few years, there can surely be no difficulty in understanding how, under appropriate conditions, a cell may, in the course of untold millions of years, give origin to the human race.—Principles of Biology. Herbert Spencer (1829–1903).

ANNUAL INTERNATIONAL PHOTOGRAPHY-IN-SCIENCE SALON

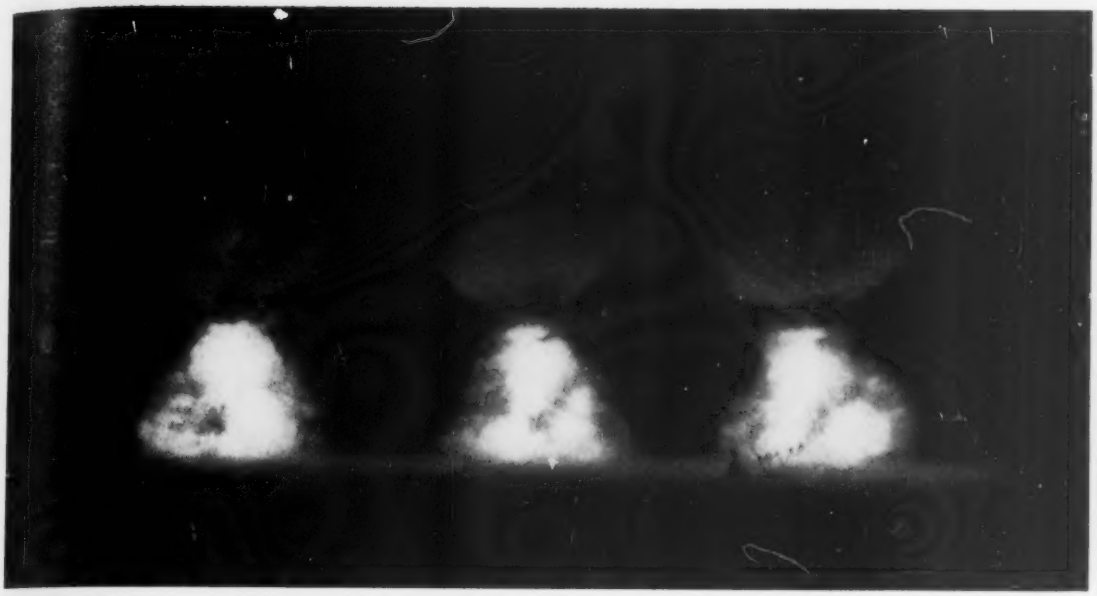


NATURAL AND ARTIFICIAL MERCURY, 5461 A

CIRCULAR INTERFERENCE FRINGES, FORMED BY FABRY-PEROT INTERFEROMETER ILLUMINATED WITH GREEN LIGHT (5461 A) FROM ELECTRODELESS DISCHARGE MERCURY LAMP: (*left*) 7 ISOTOPES OF NATURAL MERCURY; (*right*) A SINGLE ISOTOPE, MERCURY^{198}, MADE BY TRANSMUTING GOLD. LATTER PROVIDES A STANDARD OF LENGTH SUPERIOR TO ANY IN NATURE, WITH POTENTIAL ACCURACY OF ONE PART IN A BILLION. FIRST PLACE, BLACK-AND-WHITE DIVISION, 1948 PHOTOGRAPHY-IN-SCIENCE SALON. PHOTOGRAPHED BY WILLIAM F. MEGGERS, CHIEF, SPECTROSCOPY SECTION, NATIONAL BUREAU OF STANDARDS.

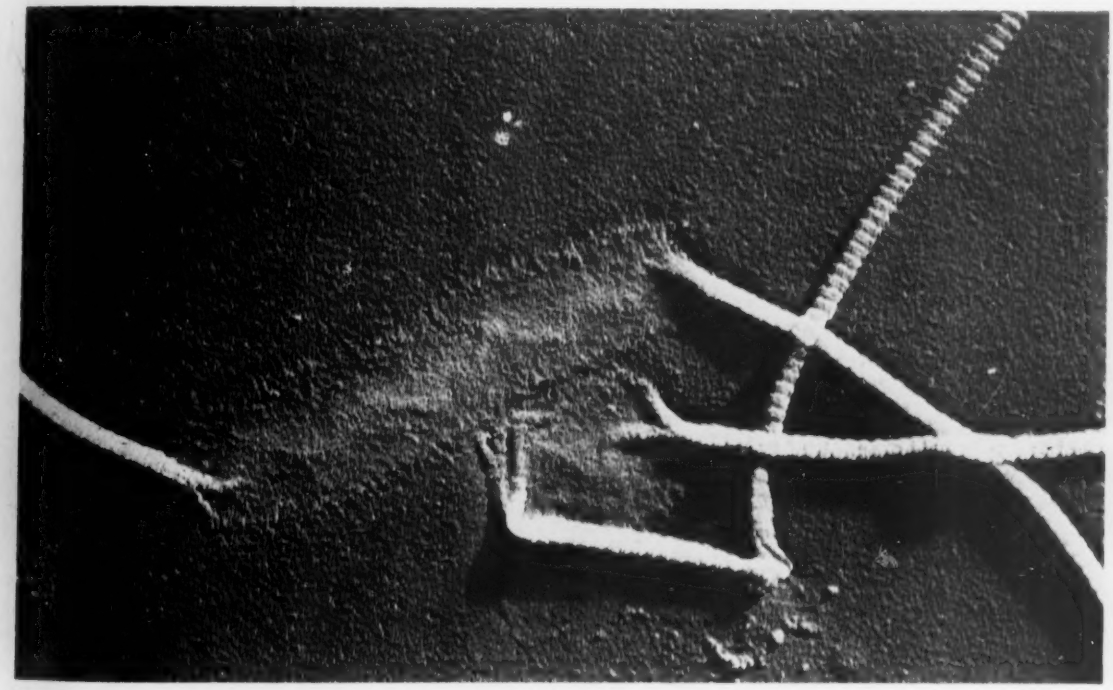
BACTERIOPHAGE LYSIS. MAGNIFICATION, 30,000 ×

ELECTROMICROGRAPH OBTAINED BY JAMES HILLIER, RCA LABORATORIES, PRINCETON, N. J., IN COLLABORATION WITH STUART MUDD AND ANDREW G. SMITH. NOTE COMPLETE ABSENCE OF DIVISION FORMS OF BACTERIOPHAGE IN LYSED CELL, BUT EVIDENCE OF PARTIALLY SYNTHESIZED BACTERIOPHAGE PARTICLES. HONORABLE MENTION, BLACK-AND-WHITE DIVISION.



EXPLOSIVE CHARGE USED IN POULTER METHOD OF SEISMIC EXPLORATION—3 CHARGES

PHOTOGRAPHIC STUDY BY THOMAS C. POULTER, STANFORD RESEARCH INSTITUTE, STANFORD, CALIF., AND WALTER K. LAWTON, SOUTHWEST RESEARCH INSTITUTE, SAN ANTONIO, TEX., WON FIRST PLACE IN THE COLOR DIVISION, 1948 PHOTOGRAPHY-IN-SCIENCE SALON. DR. POULTER ALSO WON FIRST PLACE IN THE COLOR DIVISION LAST YEAR.

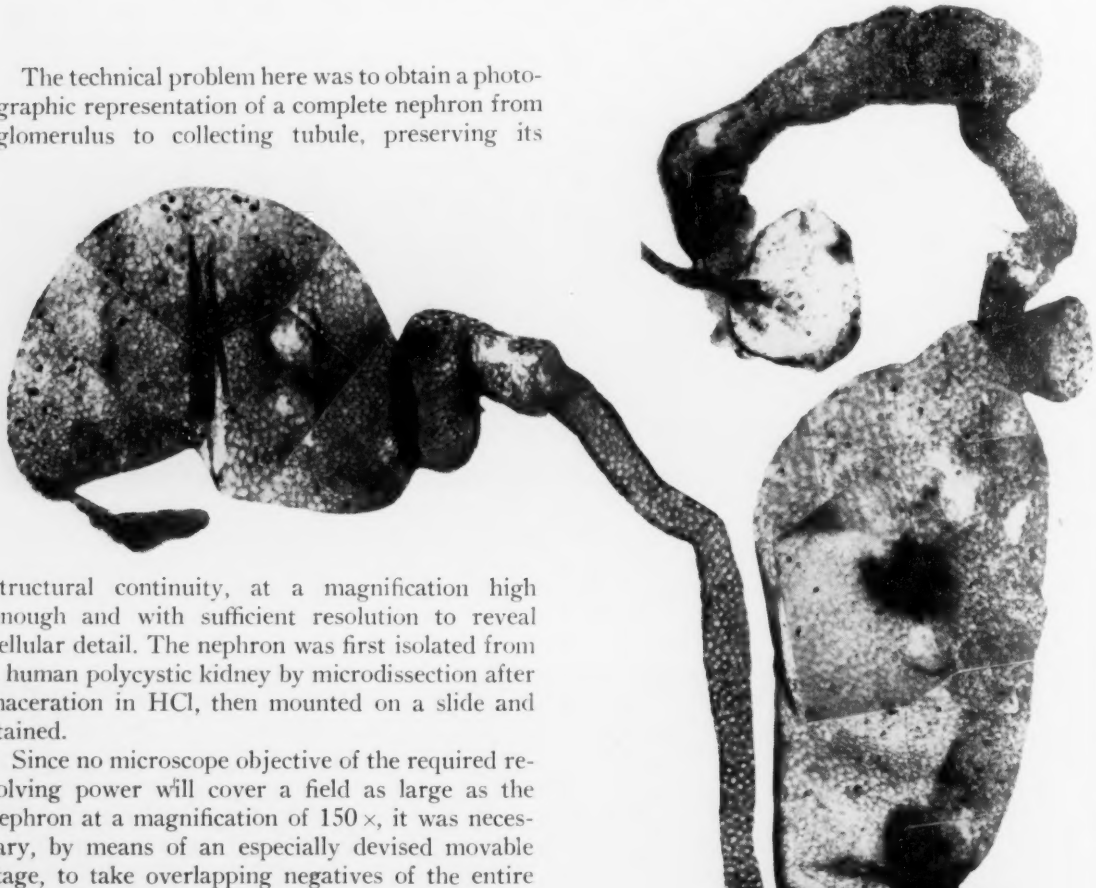


COLLAGEN FIBERS FROM COWHIDE AFTER DISINTEGRATION IN WARM WATER. MAG. 56,000 ×

SHOWING PARTIAL GELATINIZATION. (SHADOWED WITH URANIUM.) ELECTRON MICROGRAPH BY MARTIN C. BOTTY, OF THE RESEARCH LABORATORIES OF THE AMERICAN CYANAMID CO., STAMFORD, CONN.

COMPLETE NEPHRON FROM HUMAN POLYCYSTIC KIDNEY ISOLATED BY MICRODISSECTION

THIS INTERESTING PHOTOMICROGRAPH BY MURIEL C. MACDOWELL, OF THE DEPARTMENT OF PATHOLOGY, LONG ISLAND COLLEGE OF MEDICINE, BROOKLYN, NEW YORK, ALSO WON HONORABLE MENTION IN THE BLACK-AND-WHITE DIVISION.



The technical problem here was to obtain a photographic representation of a complete nephron from glomerulus to collecting tubule, preserving its structural continuity, at a magnification high enough and with sufficient resolution to reveal cellular detail. The nephron was first isolated from a human polycystic kidney by microdissection after maceration in HCl, then mounted on a slide and stained.

Since no microscope objective of the required resolving power will cover a field as large as the nephron at a magnification of 150×, it was necessary, by means of an especially devised movable stage, to take overlapping negatives of the entire object and from these, because of the lack of flatness of field inherent in microscopic optics, to select and cut out the central areas where sharp definition was preserved. Exposure, development, and printing were controlled to obtain prints of the same photographic quality; 34 selected portions were then fitted together to produce a photomicrograph of the complete nephron in its continuity from the glomerulus (*upper right*) through a cystic proximal convolution with its hyperplastic cellular proliferation down the loop of Henle and up to the cystic distal convolution, which terminates in a blind end (*upper left*).

The value of such large photomicrographs is illustrated by the problem under investigation, as the extensive area affords a view of the most important factor in the pathogenesis of the cystic disease, i.e., the sealing off of the distal convoluted tubule with disruption of the glomerulus from the resulting back pressure.

A HISTORY OF POPULATION GROWTH IN THE UNITED STATES

P. K. WHELPTON

Associate Director of the Scripps Foundation for Research in Population Problems at Miami University since 1940. Professor Whelpton has taught agriculture in New York high schools, done extension work with the U. S. Department of Agriculture, and taught farm management at Agricultural and Mechanical College of Texas.

THE increase of the population of the United States has been one of the world's outstanding demographic events. When the first census was taken in 1790 the nation had slightly less than 4,000,000 inhabitants; on July 1, 1948, it had more than 146,000,000 (Fig. 1). This represents the addition of more than 141,000,000 persons in a century and a half, or growth to more than thirty-five times the size at the beginning of the period. Few nations of the world have surpassed or even equalled this performance.

For at least 70 years (1790–1860) the population increased at a rate of 32–37 percent a decade. During this period several prominent people, among them Abraham Lincoln, made estimates of the additional growth that was to be expected. Like most of the others, Lincoln assumed the continuation of a relatively high rate of increase and forecast a population of over 250,000,000 in 1930. The 1948 forecast according to his assumptions would be 430,000,000—three times the actual population. Fortunately for our standard of living, the rate of increase declined after 1860. From 1860 to 1890 it was close to 26 percent per decade; during 1890–1900 and 1900–10, about 21 percent; and during 1910–20 and 1920–30, about 15 or 16 percent (Fig. 1). The great depression of the 1930s reduced growth to less than 8 percent from 1930 to 1940, whereas prosperity raised it to nearly 11 percent from 1940 to 1948 (which probably means over 13 percent for the current decade). During the 1950s and 1960s the downward trend in the rate of increase probably will be resumed.

Although the intercensal rate of growth fell after 1860, the number of persons added to the population during the 10 years between censuses increased to the 1920–30 period, when it was more than 17,000,000. The gain from 1930 to 1940 was less than 9,000,000, but that since 1940 exceeded 14,000,000 by July 1, 1948, and probably will exceed 17,000,000 when the 1950 census is taken. Smaller gains are to be expected during the 1950s and 1960s. By 1975 the population is likely to be between 165,000,000 and 170,000,000.

Growth from immigration and natural increase. How much of the large growth in population has come from immigration and how much from natural increase (the excess of births over deaths) is a question about which much has been written. The best information comes from a study by Joseph A. Hill, formerly Assistant Director of the Bureau of the Census. His results indicate that the white population of colonial stock increased from approximately 2,300,000 in 1780 to 41,300,000 in 1920—a gain of 39,000,000—and that the immigrants who came after 1780, together with their descendants, amounted to about 53,500,000 in 1920. In other words, approximately 42 percent of the gain of 92,500,000 in the white population between 1780 and 1920 came from the natural increase of the colonial stock, and 58 percent from the immigration which took place during the 140-year period and the excess of births over deaths in this group. It is probable that these proportions have not changed much since 1920.

Official records concerning the amount of immigration date back to 1820 and show that until 1914 the general trend was upward, with large increases occurring during periods of prosperity and smaller decreases during periods of depression. The maximum inward movement took place in the fiscal year 1906–07 when more than 1,335,000 immigrants were admitted, but 1913–14 with over 1,210,000 was a close second. The movement back to the "old country" was relatively large in the latter period, for many foreigners came to work a few years rather than to settle permanently. In consequence, the number arriving exceeded the number departing by about 770,000 during 1913–14, and probably by somewhat more in 1906–07.

Immigration was reduced sharply by World War I, but probably would have rebounded to new heights during the postwar years had it not been checked by the quota system imposed by Congres-

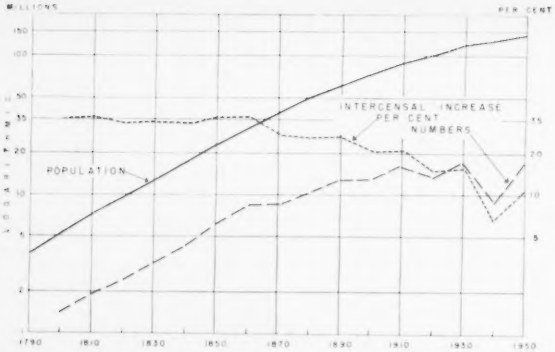


Fig. 1. Total population and intercensal increase, 1790–1950.

sional action. As it was, arrivals exceeded departures by about 3,200,000 in the 10-year period 1920–30. Administrative regulations brought further reductions and, together with the depression, led to a net *outward* movement during 1931–32. Since then the net inward movement has been increasing, and in 1947 immigrants exceeded emigrants by about 215,000. Between 1940 and 1950 the difference probably will be in the neighborhood of 1,500,000.

The net immigration during a decade was largest in comparison to the current population during 1850–60, when the excess of arrivals over departures amounted to about 12 percent of the 1850 population. A secondary high in this ratio (over 10 percent) was reached during 1880–90, but since then the trend has been sharply downward. During the present decade current immigration probably will increase the population by 1–2 percent, as it did in each decade from 1800 to 1830.

Although the excess of immigrants over emigrants has been very large in certain years, the excess of births over deaths probably has been still larger in every year, and certainly has been much larger in most years. Even in 1913, when net immigration amounted to nearly 900,000, births outnumbered deaths by about 1,200,000. Estimates prepared by Whelpton for the 10-year periods from 1800 to 1950 indicate that in most of them more than 70 percent of the current growth of the white population came from natural increase, and that in only two decades (1880–90 and 1900–10) did immigration contribute as much as 40 percent (Fig. 2). Such comparisons do not tell the whole story about the contribution of immigration to population growth, however, for a high proportion of the babies born during 1910–20 (for example) had immigrant parents.

The rate of natural increase (the birth rate minus the death rate) was largest during 1800–10, when it averaged approximately 35 per 1,000. In all except two of the following decades it has become successively smaller, decreasing to about 25 per 1,000 during 1850–60, to 13 per 1,000 during 1900–10, and reaching a low of 7 per 1,000 during 1930–40. The present decade probably will see a recovery to about 12 per 1,000. It is this decline in the rate of natural increase that is primarily responsible for the diminishing rate of population growth since 1860.

THE LONG-TIME DOWNWARD TREND OF FERTILITY

The decline in the rate of natural increase since the early 1800s reflects a more rapid fall in the birth rate than the death rate. The Birth Registration Area did not include all states until 1933, but it is possible to judge what was happening to the birth rate in earlier years by observing the ratio between the number of children and the number of women enumerated in the census. In 1800 there were approximately 1,340 children under five years of age per 1,000 white women aged twenty to forty-four, and in 1810 approximately 1,360. Since 1810 every census except one (that of 1860) has shown a smaller ratio than its predecessor. In 1860 there were 905 children per 1,000 women, a decline of about one third since 1810 (but slightly more than in 1850). In 1900 the ratio was 666, a decline of about one fourth

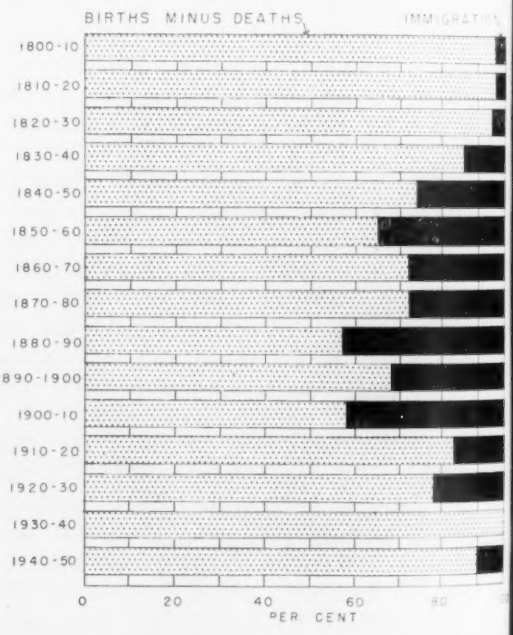


Fig. 2. Intercensal increase of the white population from births minus deaths, and from current immigration, 1800–1950.

since 1860. The last census (1940) showed a ratio of only 419, down more than one third since 1900 (Fig. 3). The next census (1950) will show a ratio substantially higher than that of 1940, and perhaps slightly above the 1930 ratio of 506. Even so, it will be only a little more than one third as large as the 1810 ratio. When an allowance is made for the great reduction which has been achieved in the death rates of small children (which should raise the ratio of children to women), it is evident that the birth rate of white women has dropped at least 65–70 percent since 1810.

Similar information about the Negro population is available from the censuses for 1850 and later years. In 1850 there were approximately 1,090 children under five per 1,000 Negro women aged twenty to forty-four. The 1880 ratio was the same, but since then there has been a decline to about 513 in 1940. The ratio will be higher in 1950 than in 1940, but probably less than 554, the 1930 figure.

Although fertility has declined throughout the United States, the reduction has progressed further in some places than in others. In general the decrease has been largest in the most urban areas and smallest in the most rural areas. Even as early as 1800 the ratio of children to white women had fallen to a much lower figure in New England (about 1,160) than in the East North Center—the area now occupied by Ohio, Indiana, Illinois, Michigan, and Wisconsin (over 1,900). Differences in 1940 may be illustrated by comparing the ratio for white women in New York City (which had fallen to 269) with that for the rural-farm population of North Dakota (which had declined only to 748). The former is barely one third of the latter. This is not an extreme difference, for several cities have a lower ratio than New York, and many rural counties have a higher ratio than the farm population of North Dakota.

Causes of lower fertility. Much has been written during past decades concerning the causes of the decline of the birth rate. Usually the blame has been placed on one or more of the following: a less favorable marriage rate, a rise in the proportion of pregnancies ending with a miscarriage or stillbirth, the more frequent resort to illegal abortion, an increase in low fecundity or sterility, and the more widespread and effective practice of contraception. The first two suggestions are of little or no importance, for the evidence now available indicates that during the past 60 years the marriage rate has become more favorable, and the proportion of pregnancies ending with a miscarriage or stillbirth has decreased. An increase

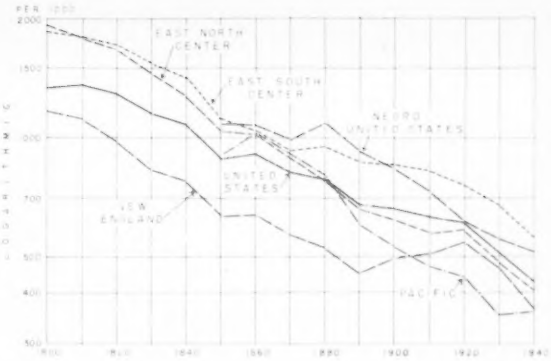


Fig. 3. Children under five per 1,000 white and Negro women aged twenty to forty-four in the United States and selected regions, census years 1800–1940.

in the practice of illegal abortion probably has been of only minor importance, for the studies that have been made indicate that not more than 5 percent of the pregnancies of urban white women, and a smaller proportion of those of rural white women, have been terminated illegally in recent years. Even if there were no illegal abortions in 1810 and the best current estimates are only half as large as they should be, less than one tenth of the decline in fertility would be accounted for.

The role of low fecundity or sterility is more obscure. There is little reliable information as to how much the birth rate is currently reduced by defects in the reproductive system, but a study of nearly 2,000 native white Protestant couples in Indianapolis indicates that it may approximate 25 percent. Still less is known about the situation in earlier decades. Some gynecologists believe that there has been a notable decrease in the physiological capacity of couples to produce children, and others that the change has been small. But even if no couples were sterile or low in fecundity in 1810 and if these conditions have reduced the birth rate by 25 percent during recent years (which appears to be an extreme combination of assumptions), only about 15 percent of the decline in fertility between 1810 and 1940 would be explained.

Much evidence is available to show that attempts to restrict family size by the practice of contraception have been widespread and effective during recent years, and have lowered the birth rate of large groups of the population by as much as 65–75 percent. That similar attempts were made early in the nineteenth century is certain. That they reduced the birth rate much less at that time than in recent years is proved by the relatively small difference that formerly existed between the actual birth rate and the maximum that is

biologically possible. It appears certain, therefore, that the birth rate has dropped 65–70 percent since 1810 primarily because couples have wanted smaller families and have made great progress in learning how to restrict the number of children by practicing contraception.

The recent rise in the birth rate. Although the birth rate declined most of the time from 1810 to 1933 and probably was less than one third as large in the latter year as in the former, it has risen substantially since 1933. In that year there were 18.4 births per 1,000 persons in the population; in 1947 there were 27.2. The increase is more striking when numbers are considered, for the 3,910,000 births of 1947 exceed by nearly 1,600,000 the number in 1933. Does this recent rise mean that the long-time downward trend of fertility has been reversed, or is it a temporary phenomenon resulting from World War II and associated conditions? To answer this question it is helpful to consider separately the rates for first births, second births, third births, and births of other orders per 1,000 women aged fifteen to forty-nine, and to deal only with native white women (who bore over 85 percent of all the children in 1947), because birth certificates are filled out more accurately for these women than for colored women.

From 1920 to 1933 the rate for first births, second births, and births of other orders declined in a rather similar fashion, the reductions varying only 26–40 percent. From 1933 to 1947, however, the trends of these rates have differed widely. At one extreme is the rate for first births, which rose in most years from 1933 to 1942 and reached a record-breaking high of over 40.0 in 1947. This is almost double the low of 22.2 in 1933, and well above the high mark of 32.6 reached after World War I. At the other extreme is the rate for eighth and higher order births, which declined in most years from 1933 to 1942, and in 1947 was about 25 percent smaller than in 1933 and 60 percent below the maximum following World War I (Fig. 4). Rates for intermediate birth orders have followed intermediate trends. Thus, the decline lasted until 1935 for second births, until 1937 for third births, until 1939 for fourth births, and until 1942 for sixth and seventh births. Similarly, the reaction from the lowest rate has been small for sixth and seventh births, and increasingly larger for fifth, fourth, third, and second births.

An analysis of these changes and of changes in other conditions during the period in question suggests certain conclusions. One is that the rates for first and second births have varied from year

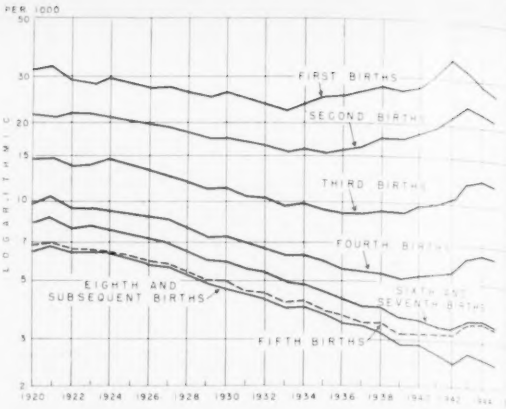


Fig. 4. First births, second births, etc., per 1,000 native white women fifteen to forty-nine years of age, 1920–45

to year primarily in accordance with prosperity and depression (modified by peace and war), but have shown no definite upward or downward trend. In other words, during the 1920–47 period as a whole, there seems to have been little change in the proportion of couples having no child or one child. A second conclusion is that since 1920 the proportion of couples having eight or more children has been declining rapidly and almost regardless of changes in economic, military, or other conditions. Conclusions regarding families of two to seven children are between these extremes.

These specific conclusions have certain corollaries. First, the rise in the birth rate from 1933 to 1943 and from 1944–5 to 1946–7 represents primarily the occurrence of births that were postponed by the depression and by World War II, respectively. Second, when these postponed births have been made up there will be a sharp drop in the birth rate and then a resumption of the long-time downward trend. The decline will be more gradual than formerly, however, because the rates for fifth and higher order births already are so low. Third, the permanent stoppage or reversal of the long-time decline will require that many couples of the type that formerly would have had no child or only one or two children decide to have an additional one or two. Only such a development can offset the further declines which are expected in the rates for fifth and higher order births.

THE DECLINE OF MORTALITY

Although mortality rates have not dropped as much as fertility rates since 1810, striking reductions have been achieved. What happened during the first half of the nineteenth century must be judged by a few studies, for deaths were not

registered in most states, the records kept by the remaining states were incomplete, and the census does not provide a measure of mortality as it does of fertility. In 1793 Edward Wigglesworth estimated that in Massachusetts four years earlier there were 27.8 deaths per 1,000 persons. In 1857 E. B. Elliott estimated the 1855 death rate as 21.4. If these estimates are correct, the Massachusetts rate was cut nearly one fourth in 66 years. Little progress was made between 1855 and 1895, the rate being between 18.0 and 22.0 in most of these years, but between 1895 and 1921 the likelihood of dying within a year was cut by more than one third. During subsequent years the Massachusetts rate has remained fairly constant between 11.1 and 13.2. Since 1933 (when the death registration area became nation-wide) the United States rate has been slightly below the rate for Massachusetts, but has been even more stable (between 10.4 and 11.6).

Death rates by age. A better appreciation of the progress made in postponing death may be obtained by considering what has happened at various ages in Massachusetts, one of the states with the best early records. In general, the proportional change in the death rate has varied from a large drop for young children to little if any decline for elderly people. During most years from 1855 to 1900, there were 190–225 deaths per 1,000 babies less than one year old. By 1920 this infant death rate had been reduced to 112; by 1946 it had been cut to 33—less than one sixth of the rate before 1900 (Fig. 5). The biggest single factor has been the control of diarrhea and enteritis, which formerly killed more than 40 babies of each 1,000 born, but which now kill fewer than 3.

Striking progress toward saving life has also been made among children aged one to four. In Massachusetts their death rate was cut from 40 per 1,000 in 1855 to 2 in 1945, a reduction of over 95 percent! After childhood, however, efforts to prevent death have been less successful age by age. At ages ten to fourteen the rate was cut by about 85 percent during the 90-year period (from over 5 to less than 1); at ages twenty to twenty-nine the relative drop was equally large (from more than 12 to less than 2); at ages forty to forty-nine the reduction was about 60 percent (from 14 to less than 6), but at most older ages the downward trend has been slight, or absent (Fig. 5). Thus at ages sixty to sixty-nine and seventy to seventy-nine the probability of dying during the year was about as great in Massachusetts in 1935 as in 1855.

Mortality rates for the nation as a whole in recent years have been similar to those for Massachusetts at ages over fifty, though slightly higher at younger ages. The latter rates have been reduced substantially since 1935, however, and there are some indications of real progress in postponing deaths among older persons. If conditions should continue as they were in 1945 the average white baby would live to age sixty-five—about twice as long as was the case 150 years ago.

Causes of death. The large reductions in death rates from infancy to middle life are the result of scientific progress, general economic betterment, and expanded public-health programs. Certain contagious or infectious diseases which were important causes of death only 50 years ago have been almost eliminated, and the mortality from others has been greatly reduced. Typhoid fever and diphtheria each killed more than 28 persons annually out of every 100,000 of the population in 1900–02, and measles, scarlet fever, and whooping cough more than 10 each, but in 1945 these five diseases combined took a toll of less than 4 (Fig. 6). The bronchitis death rate has been reduced by a larger absolute amount in the same period (from 40 to 3), and the pneumonia and influenza rate still more (from 190 to 52). The largest reduction of all is in tuberculosis mortality, the rate for this disease having been cut from 186 to 40. A very large reduction has also been achieved in the rate for diarrhea and enteritis, from 122 to 9 deaths per 100,000 persons per year.

In contrast to the foregoing is the relatively little progress made in preventing deaths from certain other causes. The rate for intracranial lesions of vascular origin (popularly called apoplexy) declined only from 106 per 100,000 persons in 1900–02 to 98 in 1945, and that for nephritis from 90 to 67 (Fig. 6). More discouraging is the fact that

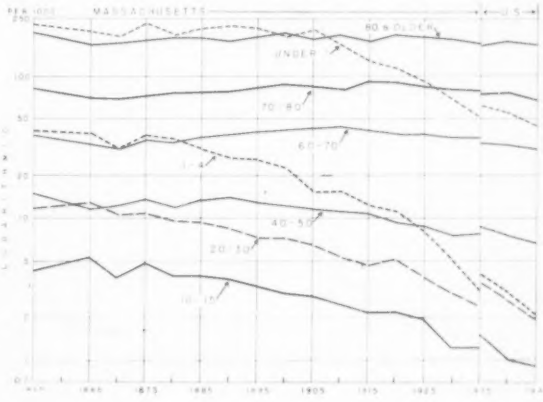


Fig. 5. Deaths per 1,000 persons of specified ages, selected years, Massachusetts, 1855–1935, and United States, 1935–45.

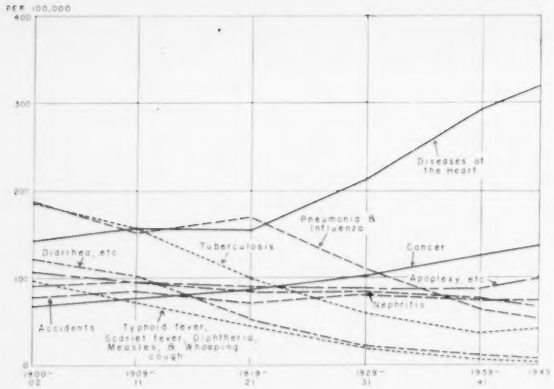


Fig. 6. Deaths from important causes per 100,000 persons in the death registration states, selected years, 1900–02 to 1945.

recent death rates for cancer and diseases of the heart (134 and 322 per 100,000, respectively) are more than twice as large as formerly; each of these is now responsible for many more deaths than any other cause. Part of the greater incidence of cancer and the so-called degenerative diseases is apparent rather than real, however, because diagnosis has become more accurate and because there has been an increase in the proportion of the population in the older age groups where these causes of death are most prevalent. But aside from such factors the increase in the rates for these causes has been sufficiently large to justify the greater attention now given to them in the national health program.

Future trends. Whether the over-all death rate will decline substantially during the next 50 years will depend in part on the continuation of the progress made in controlling such diseases as tuberculosis, influenza, and pneumonia. Here the outlook is favorable, for the more widespread utilization of present knowledge regarding the causes and control of these diseases would lower the number of deaths to a point substantially below the present levels. There is also a reasonable expectation that more effective measures for the control of these diseases will be developed, and that they will be applied to larger sections of the population as public-health programs are intensified and extended.

Although the degenerative diseases have not as yet been brought under control, there is hope for the future. Because certain damaging infectious diseases (e.g., scarlet fever, diphtheria, and typhoid fever) have almost been eliminated, a substantial reduction should occur in the organic impairments and aftereffects caused so commonly by such diseases. As these sequelae are reduced in frequency there should be a reduction in the number of organic breakdowns or a postponement of these breakdowns until later in life. Similar gains should result from the more recent campaign to control venereal disease. For this reason, and because of improved techniques for early diagnosis, there should be some reduction in the mortality from the degenerative diseases even without the discovery of better methods for their prevention or treatment. In view of the great amount of research being done on the causes and control of cancer, it is quite possible that the number of deaths from this disease will be much lower before many years pass.

MIGRATION WITHIN THE UNITED STATES

Americans have long had the reputation of being "on the move" to a greater extent than most people. Since the early 1600s the frontiers have gradually been pushed back, and the settled areas expanded from small communities on the Atlantic coast to all the habitable portions of our 3,000,000 square miles. During most of the time the main movement was westward, in accordance with Horace Greeley's famous advice, "Turn your face to the great West, and there build up a home and fortune."

Interstate migration. The first reliable information regarding the amount of internal migration comes from the census of 1850, and is based on a comparison between the number of persons born in a state and the number (excluding foreigners) living in that state. In 1850 each state from Georgia to Maine (and Kentucky and Tennessee as well) was found to have sent its native sons to other states in greater degree than it had attracted the native sons of other states. In contrast, Florida and each state west of those mentioned had attrated natives from other states in larger numbers than it had sent its own natives to other states. The net loss through out-migration had been especially large for South Carolina, the number of persons living in South Carolina being only slightly more than half as large as the number born there. Vermont, Connecticut, Virginia, and North Carolina also had suffered a net loss of 25 percent or more. California had had the largest gain, for it contained more than ten times as many people as had been born within its boundaries! Ratios of more than 2 to 1 occurred for Michigan, Wisconsin, Iowa, Arkansas, and Texas.

Although a large westward movement continued after 1850, its influence gradually was offset by an increasing movement northward and east-

ward. By 1940 this had gone so far that the Northeastern states as a group had attracted about as many people born in other regions as they had lost to other regions by the departure of their own natives (Fig. 7). The North Central region had gained somewhat in its exchange with the South and Northeast, but had lost heavily to the West. The South had lost slightly to the Northeast and more heavily to the North Center and West. The West had gained greatly from the North Center, moderately from the South, and slightly from the Northeast. Vermont and Iowa are the only states showing in 1940 a net loss of 20 percent or more through interstate migration, but 15 other states (mostly in the Midwest and Southeast) had lost more than 10 percent. California had benefited the most from migration, for nearly 3,400,000 persons—almost half its population—had been born in other states. Florida, Arizona, Nevada, Oregon, and Washington also owed much of their pre-1940 growth to a net inward movement.

During World War II the migration of the civilian population followed the general pattern just described but took place on a larger scale. Nearly 15,000,000 civilians were classified as migrants during the three years 1942–44, compared with about 14,000,000 during the five years 1935–39 and smaller numbers during preceding periods. Again it was California that attracted outsiders in greatest numbers, with some of its neighbors not far behind. In consequence, California now ranks third in population, having passed Ohio and Illinois since 1940; before 1950 it probably will pass Pennsylvania and rank second.

Migration by color, age, and sex. The white and colored populations appear to have been equally mobile during recent decades, for in 1940 approximately one out of four persons in each group was living outside the state in which he was born. The situation changed temporarily during 1935–39, when the proportion of the population classified as migrating was half again as high among white persons (12.3 percent) as among colored persons (8.5 percent). From 1940 to 1947, however, about one person in five in each group has moved across county lines.

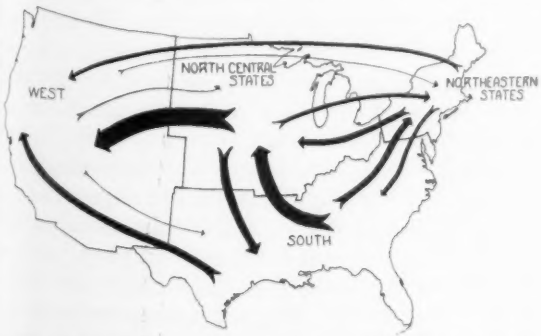


Fig. 7. Birth Residence Index: 1940. 12,792,000 migrants. (From an article in the *Amer. Soc. Rev.* by Henry S. Shryock and Hope T. Eldridge. Used by permission.)

Migration rates were highest for young adults and lowest for elderly persons during 1935–39 and 1940–47, and probably during most other years also. Because an important proportion of the young adults who move have married previously and started their families, the migration rate for children under fourteen has been higher than that for persons middle-aged or older. Youngsters of high-school age (fourteen to seventeen) have been least likely to move among the population under fifty.

Although men are usually thought to be more foot-loose than women, the proportion that makes one or more moves has been about the same for each sex. This is confirmed by a comparison of state of birth and state of residence for 1940, and also by the census data on migration during 1935–39 and 1940–47. If the number of moves could be counted, however, it might be found substantially higher for males than for females.

The rural-urban movement. For many decades a high proportion of the persons migrating within the nation were moving to new land, which they began to farm. During the nineteenth century the situation was completely changed by the industrial development, for it greatly stimulated the growth of cities. In 1790 nearly 95 percent of the population were classified as rural, and only 33,000 people lived in the largest city (New York). By 1850 the rural proportion had dropped below 85 percent, and New Yorkers numbered over 500,000. Since then the urban trend has been accelerated. By 1900 barely 60 percent of the population lived in rural areas, nearly 20 percent were in cities of 100,000 or more, and three cities were over the million mark. By 1940 only 43 percent were rural, 29 percent were in cities of 100,000 or more, and five cities were in the million class (Fig. 8).

The great urban growth came about in important degree through migration from other areas, but the net inward movement prior to 1935 can only be estimated roughly for cities with the most reliable birth statistics. Net migration to New York City probably accounted for 70–80 percent of its population increase from 1900 to 1910, and the excess of births over deaths for only 20–30 percent. (In this case, however, the majority of the migrants came from abroad rather than from other parts of the United States.) The proportions

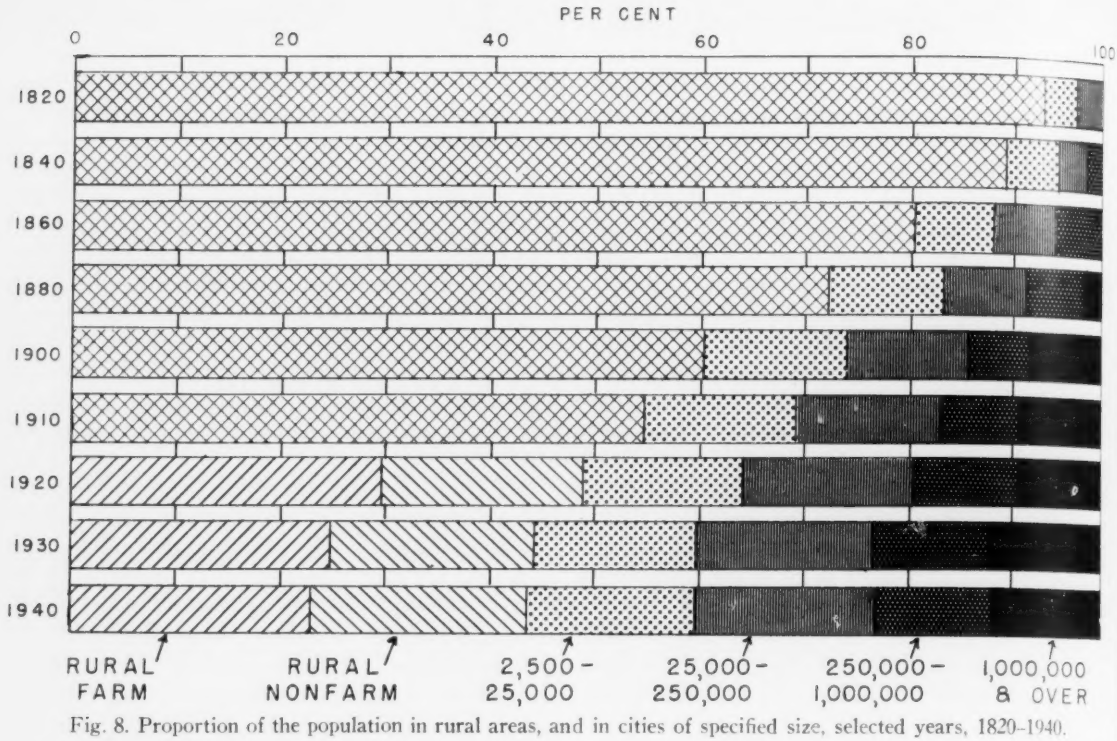


Fig. 8. Proportion of the population in rural areas, and in cities of specified size, selected years, 1820–1940.

were almost reversed during 1910–30, when about 70 percent of the growth came from natural increase and only 30 percent from migration. The two were of approximately equal importance during 1920–30 and 1930–40.

Los Angeles is the outstanding example of a large city which has grown from migration. Its population was about 577,000 in 1920; 1,238,000 in 1930; and 1,504,000 in 1940. Of the 661,000 increase during 1920–30, approximately 90 percent came from net migration and only 10 percent from an excess of births over deaths. The latter was somewhat more important during 1930–40; nevertheless, nearly 85 percent of the increase came from net migration.

The urban fringe. During the past 20–30 years automobiles and paved roads have enabled people to live in the rural areas surrounding cities, but to continue to work in the cities. At the same time the development of tractors and other laborsaving machinery has reduced not only the proportion of the population on farms but the actual *number* of persons as well, the reduction exceeding 1,000,000 from 1920 to 1940, and 3,200,000 from 1940 to 1947. In consequence, the rural population has come to consist more and more of persons who are like city folk rather than farmers. The rural-nonfarm group was less than two thirds as numerous as the farm group in 1920, but outnumbers it by a wide margin (probably about 4,000,000) at present. The movement to rural areas around cities apparently has even checked the growth of cities, for the rate of increase from 1930 to 1940 for the rural-nonfarm population (nearly 15 percent) was almost double that of the urban population (less than 8 percent). A more striking illustration of this tendency is found in the 140 metropolitan districts, where the rate of gain from 1930 to 1940 was about 30 percent for the rural-nonfarm population, but less than 6 percent for the urban population.

Decentralization? Unless war can be prevented, or at least as long as atomic bombs can be used in war, it seems criminal from a military standpoint to permit further increases in the concentration of population in and around large cities. To stop the trend in this direction, however, probably will require government action of some type. Private enterprise apparently finds it more profitable in a large proportion of cases to expand facilities for production in or near large cities than elsewhere. To sacrifice lower costs during an unknown number of years in return for making the nation somewhat stronger from a military standpoint and the factory safer in case of war is a choice few business executives can make for various reasons. So far there has been much talk of the need for real decentralization of population and industry, but

little or no progress toward bringing it about. Unless some agency capable of keeping world peace is developed, it is to be hoped that some feasible method of stimulating decentralization will be devised.

THE COMPOSITION OF THE POPULATION

Country of origin. The proportion of white persons in the population increased steadily from 1790 to 1930. In 1790 approximately 81 percent were white; by 1930 this had risen to nearly 90 percent. Since then the Negro group has gained slightly in relative importance. Others—Indians, Chinese, Japanese, etc.—have amounted to less than 1 percent of the population in each census. Indians would have been more important in the censuses before 1890, however, if those in Indian Territory and on all Indian reservations had been enumerated.

When the nation obtained its independence most of the white population was of British origin (nearly 80 percent), with the German group second (over 7 percent), and the Irish third (over 4 percent). Although many thousand Britons immigrated in subsequent decades, they were outnumbered by Germans and Irish, especially around the middle of the nineteenth century. Toward the the end of that century the immigration stream changed more radically in composition, the proportion of persons from northern and western Europe decreasing and the proportion from southern and eastern Europe rising rapidly. The net result was that by 1920 persons of all northern and western European stocks made up about 83 percent of the white population (compared with about 98 percent in 1790) and persons of all other stocks nearly 17 percent (compared with about 2 percent in 1790). The British group had declined to about 44 percent, the German group had risen to over 16 percent, the Irish group to over 11 percent, the Polish group to over 4 percent, and the Italian group to over 3 percent.

The southern and eastern European group would have become considerably more important since 1920 had it not been for the quota restrictions that have been in force since shortly after the close of World War I. During most of this period the number of immigrants from countries in the Eastern Hemisphere could not exceed 150,000 in any year, and the proportion of the 150,000 allotted to each country depended on the importance of that stock in our 1920 population. In consequence, there have been no large changes since 1920 in the relative numbers of persons belonging to the various stocks. Because the more recent immigrants had higher birth rates than others part of the time after 1920, the Italian, Polish, and other southern and eastern European stocks probably compose a slightly larger proportion of the population at present than in 1920.

Little change in the distribution of the population by country of origin is to be expected in the next few decades. Relatively few persons—barely 0.1 percent of our present population—can come annually from quota nations, and relatively few actually do come from nonquota nations. Moreover, as mentioned above, immigrants from quota nations are distributed by stock like our 1920 population. Although the millions of immigrants from southern and eastern Europe who arrived during 1900–25 had larger families than the rest of the population, and for a time contributed more than their share to population growth, most of their children have adopted the American fertility pattern and have small families.

The sex ratio. Every census has shown more males than females in the United States. The number of males per 100 females varied between 102 and 105 from 1790 to 1900, reached a high of 106 in 1910, and declined to 100.7 in 1940. Subsequent estimates indicate that the decline has continued, and that the ratio was slightly above 99 on July 1, 1948. The excess of males in the population of 1940 and earlier years was due in part to the excess among immigrants, the number of males per 100 females in this group varying between 140 and 230 in most years prior to World War I. The more important factors in the aggregate, however, were the birth of about 105 boys for every 100 girls, and death rates for females which were nearly as high as those for males. The decline in the sex ratio since 1910 has resulted from the large reduction in immigration, and from greater progress in lowering death rates of females than of males. Because the full effect of the latter has not been felt as yet, further declines in the sex ratio are to be expected.

Although males have outnumbered females in the total population until very recently, the reverse has been (and still is) true for old people. At age eighty and over there have been only 80–90 males per 100 females—the result of higher death rates for men than for women, especially after age forty. The two sexes have been equally represented among persons in the seventies, but males usually have predominated at ages under seventy. The general pattern has been as follows: Starting with about 105 males per 100 females at birth, the sex ratio has declined until the teen ages, because of higher death rates for boys than girls. Formerly

the ratio rose from the teens to middle life because of the excess of males among immigrants, going as high as 118 at ages fifty to fifty-four in 1910. In recent years, however, the decline has continued to the young adult ages and the subsequent rise has been smaller. At present there probably are slightly fewer men than women at most ages from twenty-five to forty-five, and only a small excess of men at ages fifty to sixty-five. These changes may have reduced slightly the ratio of prospective grooms to prospective brides at the ages when most persons marry for the first time, but it is unlikely that they have had, or will have, a significant effect on the proportion of men and women who marry at some time.

Age composition. As the nation has become older the average age of its population has risen also. In 1800 half the people were sixteen or younger; a century later the median age had risen to almost twenty-three years; now it is almost thirty. Future years are sure to see a further rise, perhaps to thirty-four by 1975.

The aging of the population has come about primarily because of the large decline of the birth rate, which has meant relatively fewer babies year by year. With over 50 births per 1,000 persons (as before 1830), nearly 20 percent of the population were less than five years old, but with fewer than 20 births per 1,000 persons (as during the 1930s) it is impossible for as much as 10 percent of the population to be younger than five. Second in importance during the past 30 years has been the large decrease in immigration. The millions of young adults who arrived during 1900–14 have gradually moved into the older age groups and have not been replaced in the younger groups because of the restrictions imposed on the inward movement. Third in importance during this period (but second earlier) is the large reduction in the death rates of children and young adults. This has increased the percentage of persons living to age seventy or older from less than 20 for those born in 1780 to more than 33 for those born in 1880, and will bring much larger gains in longevity for the babies of later years.

The foregoing changes have been accompanied by a substantial decrease in the relative number of children in the population. In 1820 children under five made up over 18 percent of the population; by 1900 the proportion had declined to barely 12 percent, and by 1935 to less than 8 percent (Fig. 9). The rise in the birth rate since 1935 will make this group exceed 9 percent of the population in 1950, but subsequent years are almost certain to see a further decline, probably to well below the 1935 ratio.

School-aged youngsters and their older brothers and sisters (fifteen to nineteen), most of whom are also attending school or have only recently gone to work, also have declined greatly in relative importance. In 1820 this group composed nearly 40 percent of the population; by 1900 the percentage had decreased to 32 and by 1945 to 24. Still further reductions are in prospect, though prob-

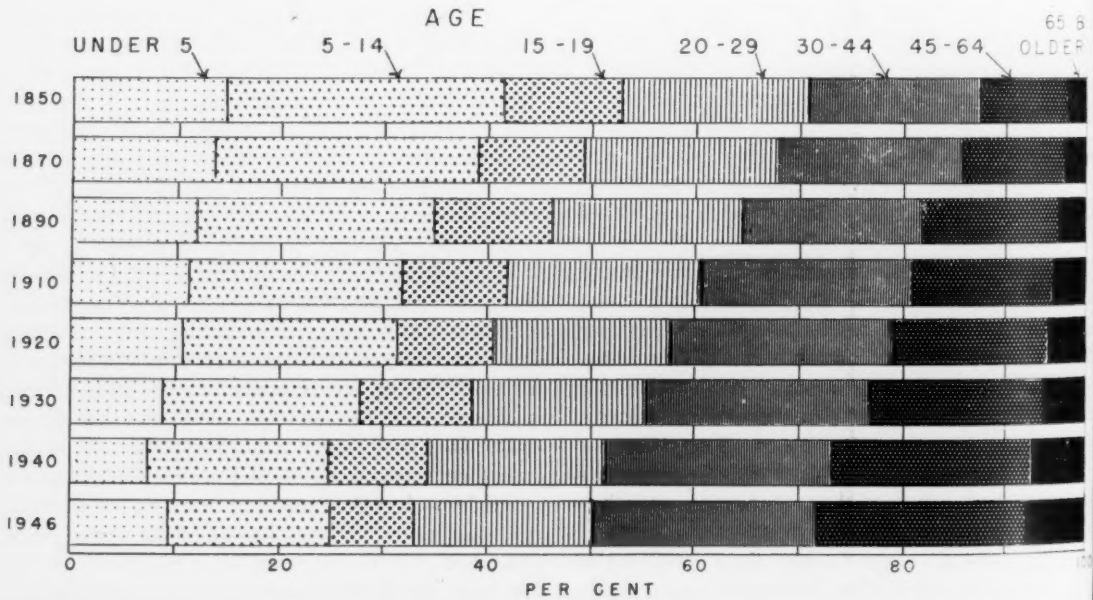


Fig. 9. Proportion of the population in specified age groups, selected years, 1850–1946.

ably not to as low as 20 percent until after 1975.

The fluctuations in the school-aged group during the next few years will cause serious problems for school administrators. In September 1947 there were approximately 2,500,000 children of the usual age to enter the first grade. In September 1953 there will be more than 3,700,000—an increase of 1,200,000, or nearly 50 percent! (Fig. 10.) The increase in the potential grammar-school population will also be striking. In September 1947 there were about 18,000,000 youngsters aged six through thirteen, but in September 1953 there will be about 23,000,000. This gain of 5,000,000 (over 27 percent) will have grade schools bulging at their doors unless there is an adequate building program during the next 5 years. Finding a sufficient number of teachers for the first grade will be very difficult in 1953 unless the teacher-training institutions turn out many more than the usual number of qualified persons during the next few years, or unless former teachers can be induced to return to classroom work.

The proportion of young adults (twenty to twenty-nine) in the population has been comparatively stable, fluctuating between 16.8 and 18.8 percent for more than a century. This has been the pivotal group, for all younger groups have had medium-to-large decreases in relative numbers, and all older groups medium-to-large increases. For ages thirty to forty-four as a whole the increase has been large, persons of these ages making up nearly 22 percent of the population in 1945, compared with less than 15 percent before 1830. Because twenty to forty-four includes the best working ages in most occupations, the increasing concentration of the population in these ages has been one of the factors tending to raise the level of living in the United States during past years. The group will have no such influence in the future, however, for it probably will change but little in relative size during the next few years and then decrease slowly.

Persons middle-aged or somewhat older (forty-five to sixty-four) are more than twice as numerous compared with other persons as they were a hundred years or more ago. In 1840 this group included about 9 percent of the population; now it includes over 20 percent (Fig. 9). Further increases are to be expected in the future, probably to between 23 and 26 percent by 1975. Nearly 90 percent of the men and 20 percent of the women of these ages were gainfully employed in 1940; they made up more than one fourth of the labor force. By 1960 this proportion is expected to rise to nearly one third. The success of employers in

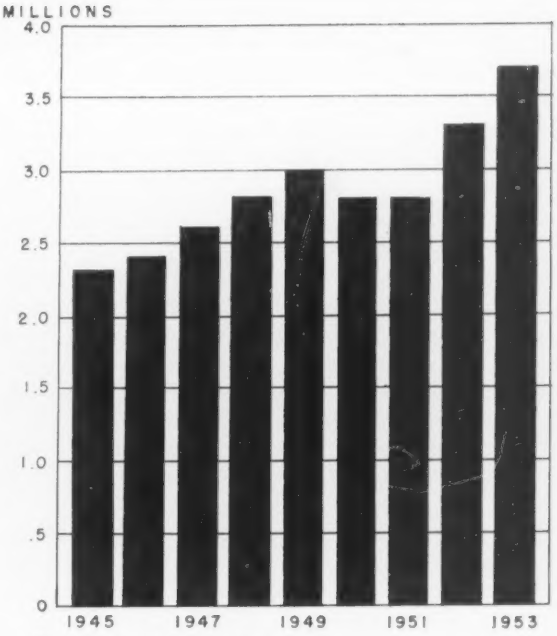


Fig. 10. Number of children of the usual age to enter the first grade, September, 1945–53.

industry and other fields in adapting their activities so as to use to advantage many more older workers and fewer younger workers may have a significant effect on the change in our level of living during coming decades.

It is the elders—persons sixty-five and over—that have had the largest relative gain in the population. In 1840 less than 3 percent of our people were sixty-five or older; now more than 7 percent are in this group. During the next 25 years the proportion probably will rise to about 11 percent—more than half again as large as at present. This large increase will call for a corresponding expansion in the services and institutions needed by older people.

Although the number of dependent elders will increase greatly in the future, the total proportion of dependent persons in the population will change little if any because of the decrease in the proportion of the population younger than eighteen. In other words, no marked increase is to be expected in the ratio of consumers to producers. Instead, the tendency for a larger proportion of women to enter the labor market is likely to lower the consumer-producer ratio.

OUTLOOK FOR THE FUTURE

The continued slowing up of population growth in the United States that is to be expected is highly desirable from the standpoint of the pres-

sure of population on resources. Moreover, as less of our productive energy needs to be devoted to providing the things needed by additional people, it should be possible to devote more of it to producing things for the people already here. The resulting change in emphasis should help to raise the standard of living.

In the old days of hand-to-hand conflict a nation with a slowly growing population gradually became weaker compared with a nation with a rapidly growing population. At present, in contrast, military might depends in large measure on capacity for economic production. This in turn depends in important degree on ability to have a high standard of living, for it is the difference between total production and production needed for subsistence that influences the amount of war matériel that can be made available. In consequence, a gradual deceleration of population growth by tending to raise our level of living, should increase our power to conduct modern warfare if it is necessary for us to do so.

Whether the expected aging of the population proves to be a serious handicap remains to be seen. Young adults are supposed to be more progressive, ready to change and adopt what is new; middle-aged and older persons are supposed to be more conservative, suspicious of change, and desirous of maintaining the *status quo*. Because of these beliefs the relative youthfulness of our population is cited frequently as one of the reasons why we have achieved the high standard of living we enjoy. If our youthfulness has had an important influence in the past, the rising proportion of older people will be a disadvantage in the future unless steps are taken to retard mental aging. Psychologists say that ability to learn reaches its peak at about age twenty, and declines extremely slowly prior to age fifty. In consequence if mental aging occurs in important degree before fifty, it is because of a decline in the use of the ability to learn rather than in the ability itself. During future years it should be possible to change the attitudes of many people toward continuing to learn, and to expand and broaden adult educational programs in sufficient degree to more than offset the adverse effect of the rising average age of the population on its ability to progress.

UNIVERSITY OF MICHIGAN LIBRARIES

ON TIME AS A PRODUCT OF MOTION

THOMSON KING

Mr. King is a graduate of Lehigh University. His article "On Life as a Separate Entity" appeared in the February 1947 issue of The Scientific Monthly.

THE LONG AND THE SHORT OF IT

The tooth of time gnaws all material things,
In vain the mountain's granite mass resists
Motions of wind and rain and heat and frost
Ever at work tear down the loftiest range.

In ceaseless flow the tireless tides of time
Sweep future through the present to the past.

Motion begets events which bring us time,
On this we build to show what time must be.
The thing that makes our universe alive
Is that which gives it time, else it were dead.
On motion look, for there you will behold
Not one, but all the qualities of time.

IF YOU ask an inteligent person for a definition of the word "time" he will usually hesitate a moment while he decides that both you and your question are quite simple and then tell you it is duration. If you consult a dictionary to be sure you understand the meaning of duration, you will find that it is "continuance of time." You are right back where you started. If you read a few of the many books that have been written about time, you will learn a great deal about the history and philosophy of the subject, from Aristotle to Einstein. You will find other definitions and many interesting facts and fancies, but you will probably end with the feeling, or at least the suspicion, that the epitome of all you have learned, the final answer to your questions, is "Time is time."

I wish to propose and briefly discuss a definition that (whatever it may be) is not circular, for it defines time in terms of something more fundamental than itself. This definition, which is also a hypothesis, is: *Time is a product (or derivative) of motion.* A theory is established when its consequences are sustained by observed facts. I hope to be able to show that many of the accepted qualities and aspects of time follow as natural consequences of the definition. Something must first be said about time in general to clarify and illustrate the definition and our approach to the subject. We shall then present and examine the evidence we can adduce to sustain the definition. If this seems satisfactory, we shall consider very briefly some of the consequences of the acceptance of the hypothesis contained in the definition.

There is an age-old argument as to whether the study and discussion of time belong to science or metaphysics. The whole structure of science is built upon order in nature and intelligence in man. Its province surely includes the investigation of all entities and phenomena to which accurate observation and measurement can be applied, particularly those that are most commonly met with in the study and description of the universe and its laws. Such a conception of science indicates time as a proper subject for scientific study. In this brief article I shall endeavor to use a scientific approach and method, avoiding as far as possible the metaphysical subtleties and philosophic terminology that are so often associated with inquiries as to the nature of time. In order to conform to the method proposed, discussion must be restricted to the entities and phenomena existing in the universe, and strictly exclude anything capable of existing outside it. Therefore, all that will be said applies to time within the universe and not to time as it might concern, or appear to, a creator of the universe, who must have existed before and outside the universe he created.

The debate about time and the literature on the subject began with the Greeks and no doubt will continue as long as men think, talk, and write. After all, time is the one thing that every creature uses all its life, and the only thing of which we all

have each day exactly the same amount, no matter who or where we are. I believe the majority of those who have written about time have thought of it as a primary or elemental thing, so my definition puts me in the minority. The philosophy and the literature of time are full of references and commentary on its subjective, objective, and perceptive aspects, of its quality of becoming, and a great deal more that is beyond the scope and unnecessary to the development of my simple hypothesis. There seems, however, to be a general agreement that essential features, qualities, or aspects of time are a flow of events; an order or sequence of occurrences in which we can always perceive a relation of before and after; and a division into three parts, past, present, and future.

To say that time is a product of motion in no way denies it the quality of reality, but does emphatically deny that it is a primary, an independent, reality. If it is a product or derivative of motion, it must be a dependent or secondary reality; motion is the fundamental, the creative, reality. Consider a room in which the only source of light is an electric bulb. If the electric current flowing through the bulb increases, the light increases; if the current decreases, there is less light; if the current falls to zero, the light disappears. The light is just as real as the current, they are both forms of energy, but the light is due to, and a derivative of, the current. We must certainly consider the current primary and the light secondary. If time were a fundamental, an absolute, we could not define it as a product of something else. We might say a good deal about its qualities and attributes, as we do of matter and energy, but we could not explain it in terms of something more fundamental than itself. We can say matter is that which possesses mass, but we can only say mass is an attribute of matter. The definition of true fundamentals is always circular.

Neither does my definition imply that we should stop talking about time and substitute for that good word some reference to its derivation from motion. The definition will, I hope, lead to a better conception and understanding of the nature of time. It does imply that the time of any interval, a minute or a century, is the result of the summation, the integral of the motions that took place during the period. It is much simpler and more convenient, however, to use one word rather than many. It is better to say "Here is a dollar" than "Here is a piece of paper which the government of the United States has declared legal tender with the value of one dollar." The longer expression is useful to explain the real nature of the thing spoken of, but when this is once grasped it is natural to replace the explanation of the derivation of value by a single word.

Scientists continually use the concept of time and measure time intervals with great accuracy, just as they measure mass of matter, dimensions of space, and quantity and intensity of energy. Scientists can describe the phenomena of what Dunne calls "elementary undefinables," but they do not attempt to explain the existence of whatever they accept as a fundamental reality. If a thing is a product or derivative of something else, it can be defined and treated as such, and its existence can be explained in terms of something more fundamental than itself. To show that time is a product or derivative of motion, it will be necessary to prove that motion can create time. I think it will appear self-evident that time of itself cannot create motion.

Some philosophers have taught that time is only an abstract mental concept, having no physical basis of reality, that exists only because of our consciousness of its passage. This view, making the concept and definition of time dependent upon human perception, seems to me narrow, untenable, and useless. It has always led to unsurmountable contradictions or a resort to mystical interpretations. If time has no physical existence, what do we mean when we say that the dimensions of the unit of action, which is an invariant, a constant of the physical world, are energy multiplied by time? It is true that we are conscious of the passage of time because we are conscious of events. The perception of events is necessary to intelligent beings, but intelligent beings are not necessary to events. Events and, consequently, time existed long before there was life on this planet and would continue to exist if all life were destroyed. Mountains rise and are worn down, rivers flow, cycles and seasons, night and day, exist and recur whether or not we are here to observe them. All these changes are produced by motion. I think that a strict analysis of our perception of time will always prove that it is derived from a perception of motion. The ultimate argument for this is, of course, that if there were no motion, there could be no change, no events, no flow of time, nor perception to take note of them or their absence.

The full meaning of my definition may be most easily realized by considering what would happen to time if all motion in the universe were reduced to zero. As the rapidity of movements were everywhere reduced, events would occur more slowly compared with their present rate, but not as compared with each other; for I am assuming that all

UNIVERSITY OF MICHIGAN LIBRARIES

motions would be reduced at the same rate. A "year" would be longer than our year, but a clock would tick the same number of times, and our hearts would make the same number of beats as in normal years. Light would travel the same distance in a second, for seconds would be longer. When all motion was reduced to zero, time would disappear. When the galaxies and the molecules ceased to move, when electrons no longer spun, when radiation, losing both frequency and velocity, ceased to be energy, nothing could happen, there would be no change, no events. There could be no flow of time; everything would be frozen in an eternal present. If time were a true fundamental, we could not, even in imagination, make it disappear or cease to be, by reducing something else to zero.

To support the hypothesis that motion is more fundamental than time, it will be necessary to meet the argument that motion is a secondary reality because there is always something that moves, and, therefore, movement may be regarded as an attribute of matter, like color or density, rather than a fundamental something in itself. Since it will also be claimed by some that motion must take place *in* time, this must also be disposed of. The idea of the primacy of motion has been developing for a long time. Democritus held that all change could ultimately be resolved into movement. The greatest of the Greek philosophers debated these matters, with the general conclusion that the two great fundamentals of the universe were stuff (matter) and motion. Today the evidence for motion is far more conclusive. If we know anything of the nature of the universe, it is that it is dynamic. So far as we know, all matter in the universe, from electrons to galaxies, is in motion. We appeal to motion to establish the reality of space, for motion consists of something having gone from where it was to where it was not, which would hardly be possible unless what we call space existed—or perhaps motion creates space.

When we consider energy, we find that it is always associated with motion. Perhaps we can say *energy is motion.* The measure of the kinetic energy of matter is the square of its velocity. The energy equivalent of a given quantity of matter is the product of its mass by the square of the velocity of light. The measure of every photon of radiant energy is its frequency, which gives the rate of repetitive cycles involving some form of motion. The development of relativity theory has shown that many things are relative to motion, whereas motion seems to be relative only to other motions. Even mass varies with motion, indeed the whole c. g. system seems to be relative to motion, although of course such effects are observable only at velocities that are an appreciable part of the velocity of light. Motion seems unique in having a fixed upper limit in the speed of light, which is a constant of nature.

It will prove interesting to think a little more about a universe in which there is no motion. What can be said about a universe that contains matter, but has never had motion? By definition this universe contains nothing but matter. There can be no change, no energy, no action (energy × time), no temperature, no life, and no time. We can only describe it as an aggregation of matter, and its history as timeless existence. But these descriptions could not be made from within the universe; they could only be made from the outside by someone who has knowledge of it, or who imagines it. This person must be familiar with the qualities we have referred to as not present in the motionless universe.

I have asserted that time would not exist in the universe I have asked you to imagine. The majority who believe that time exists as a flow or sequence of events, and must have a present, past, and future will agree, but there may be some who will say, "You admit that such a universe has existence, therefore it must have time to exist in." I reply that time may be duration, but cannot be mere existence. Duration always requires the perception of before and after, but to have a before and after we must have change. Duration must have a period of time, and the permanence of something that endures contrasted with the flow of time or the change of something else. In our motionless universe there can be no periods or intervals, and no change. If we reduce the meaning of time to mere changeless existence, we strip it of every quality by which we can apprehend it or perceive its reality. Like the Cheshire cat in *Alice in Wonderland,* it shrinks till there is nothing left but the grin, and finally that disappears. Definition as mere existence makes the whole concept of time meaningless and perfectly useless.

In this motionless universe that contains only matter the concept of space, as we understand it, would be invalid and without meaning. There is no room here for the development of this idea, but the argument would be along the same lines as that for the nonexistence of time. It also seems evident that there could be no organization of matter into atoms, and consequently no differentiation of matter into elements. We would have the rest-mass of the ultimate particles without form or organization.

Let us do one more thing to our imaginary uni-

verse. We introduce the various forms of motion that exist in our real universe. There will probably be some form of explosion such as Le Maitre supposed to have occurred in the giant concentrated atom containing all matter which exploded to create an expanding universe. As soon as motion is introduced we have the kind of universe in which we live. It is endowed with energy, time, space, action (including Planck's constant), temperature, atoms, a diversity of elements, and the possibility of life. I think we can safely say the introduction of motion created both the space and time that are aspects of the universe as we know it. Perhaps this will account for the necessity of the space-time concept in the mathematical structure of relativity theory.

Minkowski said, "From henceforth space by itself and time by itself are doomed to fade away into mere shadows, and only a kind of a union of the two will preserve an independent reality." I am a little doubtful that the union of two shadows by a hyphen, or even a mathematical formula, can create an independent reality. If, however, both space and time are created by, and dependent upon, motion, we can understand why Minkowski calls them shadows, and also it is natural that they should be associated to form the four-dimensional continuum of relativity. If we can imagine each electron, or other elementary particle, as creating its own space and time we can immediately understand the reason for Pauli's exclusion principle.

I believe the foregoing illustration will be easily followed and that the birth of time will appear as a natural and inevitable consequence of the introduction of motion into a static universe. There are many other considerations that support the primacy of motion and indicate time as a derivative or product. Long ago Aristotle realized the intimate connection between time and motion when he said: "Time is measured by movement, and movement by time." This does not imply equality in primacy. The amount of metal deposited in an electroplating operation is a measure of the electric current, and the amount of the electric current is a measure of the metal deposited, but there is no question as to which is the primary cause of the operation. Einstein has shown that there is no such thing as absolute time. I think it is generally accepted that time is relative to motion. If time be defined as a derivative of motion, the fact that the rate at which clocks run (as ascertained by light signals) will be different for observers in different frames of reference is easily understood. These observed and accepted facts of relativity are a natural and inevitable result of the definition of time as a product or derivative of motion. I do not know of any other definition of time of which this may be said.

All who write about time must deal with its aspect as an order or sequence possessing three parts, and with its attribute of "becoming." The future is always becoming, or flowing into, the present, the present into the past. In fact, the true present must be thought of as a line or plane, with no thickness. It is simply the boundary between the future and the past. The philosophers have been forced to invent the "specious present," which they define (not too successfully) as the shortest time in which the mind can entertain a single thought and realize duration by experiencing the sensation of before and after; or, according to Webster: "The time span of immediate consciousness; that interval in which what is earlier may be distinguished from what is later, though both are directly present to consciousness." This was necessary to give our thoughts and actions a period with duration in which they may occur. If time were a fundamental thing in itself, we should have great difficulty in saying in what part of it our thoughts and actions occur. They certainly do not occur in the future. Since the "specious present" is obviously a convenient fiction, and the true present has no duration, there is no room for them in the present. They cannot occur in the past. We shall return to this paradox later.

At this point it is interesting to note that all events are in the past before we are aware of them. Suppose the lead of your pencil breaks while you are writing, you may observe the event by seeing or feeling it happen. It requires time for the light to reach your eye, although the distance may be only a few inches, and it takes time for the transmission of the message along the nerves from hand to brain. What was the future, in no time at all becomes the past. An event can take place in a year, a day, or a second, but the only way I can imagine it taking place in the present is by supposing it to move with the present. All events are produced and brought to us by motion. Time is the product, not the cause, of events.

The scientific approach to any subject always involves measurement; measurement requires units. In trying to define time, it will be helpful to consider its measurement and units. Measurement is always accomplished by comparing the thing to be measured with some standard or unit of its own kind; length with a standard length, weight with a unit weight. Our units of time always refer to motion, and, since motion takes place in (or even possibly creates) space, there is usually, but not

UNIVERSITY OF MICHIGAN LIBRARIES

necessarily, a spatial dimension. The year involves motion and distance; the latter does not enter into the half-life of a radioactive element, which could be used for the measurement of time. If time is truly a product of motion, it is quite natural that its measurements and units must be referred to motion.

Time intervals, like lengths, are most conveniently measured by comparison with a standard, but the measurement cannot be accomplished by direct application of the standard, as in the measurement of length. A standard of length or weight can be preserved and used many times, but time intervals, and the motions that define them, are transient things—as soon as completed they are gone beyond recall. All that can be done is to select some natural phenomenon or to contrive some machine, such as a pendulum, that will continue to repeat a cycle of movement with the least possible variation. The regular repetition of the cycle measures time and gives us our standard.

The measurement of time always involves motion, but it is possible to measure time without recourse to a repetitive cycle. We could measure a second in many ways: for instance, the time it took a weight to fall a certain distance at the surface of the earth; we could measure an hour by a sun dial or the ancient water clock—but we would always come back to motion. Even the decay of radioactive elements is brought about by the motion of particles or photons from within the atoms to the outside. I think Bolton in his *Time Measurement* has given us the final essence of the matter of units and measurement of time if we omit the word "repeated" and quote "All measurement of time depends upon uniformity or invariance of (repeated) motion."

Almost all who write about time sooner or later tell us what Kant and Bergson thought about it. This is indeed indispensable in any general and complete consideration of the subject, but the scope of this article forbids any such excursion no matter how attractive. What Newton thought and wrote is interesting and worth quoting. He realized and acknowledged the dependence of what he considered "common time" on motion, but he believed there was another kind of time that was absolute and independent of all else. For nearly two centuries the prestige of his name helped to sustain belief in the existence of time that was absolute and independent. Today the general acceptance of the theory of relativity has displaced his idea of an absolute time:

> Absolute, true, and mathematical time of itself, and from its own nature flows equably without regard to anything external, and by another name is called duration; relative, apparent and common time, is some sensible and external, whether accurate or unequable measure of duration by the means of motion and this, which is commonly used instead of true time such as an hour, a day, a month, a year.—*Principia Mathematica*, Book I, Scholium I.

Unfortunately Newton does not tell us how "absolute, true, and mathematical time" can be known, measured, or distinguished from "relative, apparent and common time," nor has anyone else. It is very possible that he was influenced by his religious convictions in his belief in the necessity for an absolute time. We also know that he thought there might be a point or region in the universe at absolute rest. If this were the case it would afford a point of reference from which absolute motion could be reckoned and, consequently, an absolute time. Today we see no hope of finding such a point; therefore, most things are relative.

From the weight of the evidence I think we may reach the following conclusions: Motion cannot be an incidental attribute of matter, rather it is motion that possesses matter. Motion gives the universe its most important characteristic as a dynamic system in which change is continuous, a sense of before and after is possible, and there is a continuous flow of the future through the present to become the past. All events are due to change, change is always the result of motion, time is a sequence of events, motion is the begetter of change, events, and time. Time is an aspect of the universe that is due to motion. A dynamic universe must have time; a static universe cannot have time.

I have already anticipated one of the most important consequences of my hypothesis by stating that the relativity of time (and the simultaneity of events) to the motion of the observer, which was demonstrated by Einstein, follows as a natural and direct consequence from the definition of time as a derivative of motion. The time of each observer is only another name for the flow of events produced by the motions of and in his frame of reference. Therefore the observations concerning time by an observer in another frame of reference with a different motion will not be the same.

I also believe that the acceptance of the definition may be useful in clarifying the relation of the past and future to the present and the question of their reality. Equally it may help in understanding and disposing of a paradox that has been a most fruitful source of perplexity and the cause of various and lengthy dissertations and explanations. I have no intention or desire to discuss either the meaning of reality or the paradox itself from the philosoph-

ical point of view, nor even to imply that I, or anyone else, can give an entirely satisfactory definition of reality. I believe the subject still stands about where Plato left it. But practically every intelligent person has his own idea of the meaning of reality, and I shall be content if he will apply it to what I have to say.

So long as time is considered as a primary and independent something-in-its-self, apparently coming from nowhere and vanishing into nothingness, there must be a strong question as to the reality of the future and the past. It is self-evident that our actions and events cannot take place in the future or in the past. This leaves us only the present, but the present is merely a boundary between the future and the past; it has no thickness, no duration in which events can occur. Here is our paradox: Events cannot take place in the true present, and they cannot take place outside it. Yet they certainly take place. No doubt this delighted Zeno, and it has continued to delight his philosophical heirs and assigns.

Let us approach this problem sticking closely to our definition of time as a product of motion and remembering to substitute motion for time in our thinking of the relation of past, present, and future. I think we shall find the whole problem has been simplified somewhat according to the following argument. I take it that the vast majority of persons (who think at all of such things) consider matter and energy as real and primary entities or, since they are at least to some extent interchangeable, two aspects of one entity. The conservation laws tell us that matter and energy are enduring realities. They change their form, but they are not destroyed. Something cannot become nothing, nor can nothing become something. If matter and energy are real, the motions of matter and energy are real. Motion is inextricably associated with energy.

All the events that we can observe, with the possible exception of certain mental and physical phenomena, are due to the motion of matter and energy. We know of time only by a succession of events, by our sense of before and after. The flow of time is the result of the succession of innumerable events produced by the innumerable motions taking place in the universe. The transition of the motions-that-are-to-be, to the motions-that-are, to become the motions-that-were, gives us our flow of the future through the present into the past. We are equally conscious of motions-to-be and motions-that-were. We recognize the reality and validity of motions-to-be every day of our lives, particularly when we are driving a car. The present is created, brought to us, and carried forward by motion. It is indeed a boundary line between the future and the past, but it is a moving line, generated and swept forward by the same agency that creates events and time, and moving at the same rate. From and because of motion the present moves, events take place, and time is apprehended by conscious beings. Since all are generated by the same motions, all develop at the same rate, and within the line (or plane) of the present. A homely analogy may help to illustrate my meaning. Consider water rising or sinking in a vessel. We can think of the surface as a boundary plane of no thickness between air and water. Yet the events of emergence or submergence of the sides of the vessel all take place in the moving plane. As the water rises, in this parable, the unsubmerged area is the future, the surface of the water is the present, the submerged area is the past.

I think we must conclude the past possesses reality because it consists of events created by motions that were real. The water that ran over the dam yesterday was real, its motions were real. Neither has become unreal because we refer to them in the past tense. The effect of past motions, at least to some extent, determines and creates the motions and events of the present and future. The motions that led to the formation of the earth and the motions of the winds and water of yesterday will certainly continue to produce events today and tomorrow. By similar reasoning the future can be assigned a potential reality. At least some of the motions that will produce the events of the future have been assured by the motions of the past. This potential reality is created and imposed by the nature of the universe and the constancy of its laws. It is illustrated whenever we say, "The sun will rise tomorrow;" or whenever an astronomer predicts an eclipse in 1960, and the event happens as predicted. This is by no means an argument for complete causality. The principle of indeterminacy seems to impose definite limitations as to the events that can be predicted, and we know too little about the nature of life and its potentialities to say much about the prediction of events in which it plays an active part. Nevertheless, unless the future has at least a potential existence, I do not see how any events could be predicted with the certainty that we constantly observe. In so far as we can predict future motions, we can predict future events.

There has been much debate as to whether time is reversible. The general conclusion seems to have been that it is not. The second law of thermodynamics, which is a statistical-probability law, predicts the inevitable increase of entropy in the

UNIVERSITY OF MICHIGAN LIBRARIES

verse, or any isolated system where work is being done, with the passage of time. This indicates the direction of the flow of time and forbids its reversal. The behavior of energy is always to run downhill to lower frequency numbers whenever it does work. Like Humpty Dumpty when he fell from the wall, the universe can never return to a former state; or, if you prefer, the probability against such a return is practically infinite.

By definition, time is the name we use to denote the integrated result of the practically infinite number of motions taking place in the universe. Most of these motions, so far as we have knowledge of them, are random. The probability that all the motions of even a very small period will be reversed in the next like period, to restore a former state, is practically zero; it may be that the nature of energy makes it not only vastly improbable, but absolutely impossible. Whether this be so or not, the probability against the reversal of all motions is large enough to effectively settle the matter. The irreversibility of time is a direct consequence of its definition. In fact, it is hard to see how time can be reversed for even a very small isolated system. Take a small clock pendulum: at the end of each double swing it returns to the same position with regard to the objects immediately around it and, so far as outward, superficial appearances go, to its former state at the beginning of the cycle. But during the cycle there have been billions of motions and changes in relation among the billions of molecules, atoms, protons, neutrons, and electrons of which it is composed. The probability that all these have been reversed to restore its original state is practically zero. As a single mass, it repeats a cycle; but as an aggregation of particles its behavior is quite different. For it, as for the rest of the universe, time can never return, a past state can never be regained.

Let us see how we would go about stopping and reversing time for one particle of matter. We will consider we have accomplished the first when no events affecting the particle have occurred for one second. Reversibility will be established when we return the particle, at the end of one second, or some other interval, to exactly the same state and conditions that existed with regard to it at the beginning. For our experiment we select a neutron, shielding it from cosmic and other radiations. The effects of molecular motion may be eliminated by reducing the particle and its surroundings to absolute zero. Perhaps it is spinning; each revolution is an event, so it will be best to stop the spin. Have we had an uneventful second for our neutron? Change of position is an event. It will be necessary to stop the rotation of the earth, its motion in its orbit, and all the motions of the solar system and of our galaxy. To restore the former state these motions must all be reversed. Even this will not suffice, for, unless all the motions of the stars within our galaxy and of the other galaxies are stopped and reversed, the position of our neutron with regard to them and their attraction for its mass continue to change, events are occurring, and the probability of a return to a former state is practically zero.

The conclusions to be drawn from the foregoing argument seem to me quite definite. The irreversibility of time for the universe, or even for one particle of matter, can be established quite simply, without appeal to the second law of thermodynamics and the increase of entropy. Furthermore, to stop or reverse time for one particle of matter it would be necessary to stop or reverse all the motion in the universe. The time of the universe is the result of all its motions, of its nature as a dynamic system. If we allow one particle of matter to have motion, we have change of position relative to that particle and a change of gravitational attraction (or, if you prefer, a change of curvature of space) for every other particle; we have change, events, and time. There is an element of unity in this view of time, motion, and the contents of the universe—it seems to be a case of one for all and all for one.

One of the interesting consequences of our definition is that we find in it an answer to the question, To what extent can life change future time? The answer is a definite yes, in so far as it can change motions-to-be, life can change events-to-be, which is changing time. Motion is the outstanding characteristic of the universe. Purposeful directed motion is characteristic of life. "We are observers who possess the power to interfere with what we are observing." Consider a river and its valley. The continued flow of the river and many events in the valley, with regard to both living and nonliving things, may be foretold with considerable accuracy provided there is no interference by man. Many events have taken place in this river and valley in the past; our experience as observers indicates what may be expected in the future if we do not interfere with the natural processes. Then men build a dam, a lake is formed, certain areas are flooded, submerged areas become dry, others are irrigated. Motion, life, and events in the river and valley are quite different because of the decision and actions of certain men. Events and their order and sequence, of which time consists, were changed in the illustration, and are being changed

every day in larger and smaller ways by the action of life. Time is certainly changed, unless someone can prove that it was time that compelled the men to build the dam. To do this one must endow time with qualities I am sure it does not possess.

Our definition also answers the question, When did time begin? The answer is, of course, When motion began. And we may add, It will end when motion ceases. These answers may not be entirely satisfactory for those who desire an actual date, but they are the best we shall have for a long time. If a way of determining the age of atoms analogous to the method now in use for determining the age of uranium-bearing rocks from their lead isotope content is ever found, perhaps we may have an indication of the age of the universe in its present form.

I think we can push our analysis of the evidence still further to disprove the argument that time is elemental and that motion takes place in time. If time were a fundamental, rather than a secondary characteristic of the universe, and not dependent upon motion, it would exist in a motionless universe and would not lose or change its nature with the introduction of motion. This absolute time that could exist without motion could have no flow, it could contain no change, no events, no past, and no future. The only quality we can ascribe to it is changeless existence. Motion could never take place in this kind of time, nor can we ascribe to it any genuine reality. It is simply a name, signifying nothing. A universe of which such time were a fundamental characteristic could never be anything but dead mass, it could contain no energy, nothing could ever happen in it. It is certainly not our universe.

In contrast to the difficulties and contradictions that infect and plague the systems that consider time as a primary reality, a something in itself, its definition as a product of motion gives it a reality that is consistent with the universe as we know it, and explains the qualities and characteristics that we constantly observe. Its flow, its quality of becoming, its past, present, and future, all that we have referred to earlier, and much more, demand motion as the creator of time. I believe that further exploration of the consequences of our definition may yield many interesting and useful advances in our understanding of time and the physical world. The idea has never been carried to its ultimate logical conclusions.

A few parallel or paired statements will serve to summarize the evidence and conclusions and illustrate the idea that all that can be said about our perception and experience of time can be asserted with equal propriety and validity if we accept its definition as a product of motion. Time is continuous, not atomic; so is motion. Time is irreversible; the motions that produce events, even for one particle of matter, are irreversible. Time is known by events; motion produces events. In time we have past, present, and future; in motion we have motions-that-were, motions-that-are, and motions-to-be. Entropy increases with the progress of time; entropy increases with the progress of motion. The dynamic character of the universe makes time possible; the dynamic character of the universe is due to motion, *is* motion. The tooth of time gnaws all material things, changing their form; the change of form of all material things is due to motion. Time is different for observers in different frames of reference; the definition of difference for frames of reference is a matter of motions. The state in which we find the universe (less than maximum entropy, less than complete equipartition of energy, and other matters) indicates that the time that has elapsed since its present state began is finite; the same assertion can be made with regard to motion.

These are but a few of many assertions that could be found to support my arguments. I have searched faithfully to find a single statement that is incompatible with the simple hypothesis I have proposed. So far my search has been unsuccessful.

UNIVERSITY OF MICHIGAN LIBRARIES

THREE METHODS OF VOTING

WILLIAM S. TAYLOR

Professor Taylor received his doctorate in philosophy at Harvard in 1921, and since 1926 has been professor of psychology at Smith College. He has published two books and various monographs and articles in dynamic and abnormal psychology and several articles in social and political psychology.

MANY groups have learned that they can reach decisions better, and can reach better decisions, through voting than through fighting. Such groups include political units, social organizations, vocational associations, boards, and committees.

Many of the same groups do not realize, however, that scarcely less important than voting is the method of voting. Often, given a variety of candidates, a wrong method of voting elects only the candidate or candidates whom the majority do not want. Sometimes a wrong method elects candidates whom not even an important minority want. Often a wrong method elects only the candidates of an organized or lucky minority. When several candidates are to be elected to a representative body, a wrong method fails to elect candidates to represent the majority, if there is one, and the minority or minorities in proportion to their numbers. Such unsatisfactory elections discourage good candidates from running for office, and discourage voters from voting conscientiously. Such elections also encourage schemers to take advantage of the poor method, to use it for selfish ends. Thus true progress is hampered or thwarted, and even the principle of voting is impugned.

It would seem that every electorate, whether large or small, should know how to choose and apply whatever method of voting is best for any given election. In what follows, we shall consider three general methods, which we shall call Ordinary Vote, Majority Preference, and Proportional Representation, respectively. (Here Proportional Representation means not the European methods of block or party listing but individual use of a single transferable ballot.) We shall find that these three methods often yield significantly different results; that each method is best for particular, everyday circumstances; and that each method is workable, not only for large elections but also, in one form at least, for elections by small associations, departments, and committees.

In Ordinary Vote, as everyone knows, the voter marks his choices either by writing names or by crosses; and election is by the largest numbers of names or crosses.

In both the other methods, the voter either arranges the names, or marks the several candidates 1, 2, 3, etc., from his first to his last choice. The ballots, however, are counted differently for each method. We shall refer to the rules for the counts later. Here let us note only that the rules are simple for Majority Preference, and fairly simple for Proportional Representation; and that they cover uncompleted ballots, erroneous ballots, the counts of the valid ballots, and various ties, even for elections by only three voters.

Majority Preference elects the candidate or candidates whom the whole electorate taken together prefers over every other candidate.

Proportional Representation recognizes that many an electorate divides naturally into subgroups, of which one may be a majority; and Proportional Representation elects those candidates who, according to the ballots, best represent the subgroups in proportion to their numbers. For example, if six conservative and three moderate voters are to elect three out of several conservative and moderate candidates by Proportional Representation, two conservatives and one moderate will be elected.

In some elections, depending upon the assortment of candidates, voters, and issues or points of view, the different methods of voting will yield identical results. Often, however, the different methods will yield very different results. Let us compare the several methods through simple examples in which the results will be different (Table 1).

In the first pair of examples, as the underlined scores show, Ordinary Vote elects Candidate C. In this case, C is favored by only three of the seven voters. Moreover, as appears when the same seven voters vote for the same three candidates by Majority Preference, C is just the candidate whom every one of the other four voters wants least. Majority Preference, on the other hand, elects the second M, the candidate who re-

ceived the smallest sum of preferences. (If he were the first choice of every voter, his sum would be 7.) He is the candidate whom the electorate as a whole wants most.

In the second pair of examples, Ordinary Vote elects the two R's. These candidates represent only three of the eleven voters. Also, as appears when the same eleven voters vote for the same lot of candidates by Majority Preference, the two R's are the candidates six of the other eight voters want least. M. P. elects the first and fourth M's, the two candidates the electorate as a whole wants most.

Nevertheless, these two candidates do not best represent the whole electorate, when we consider that every electorate reduces to individual voters, and that, in most elections, the voters arrange themselves into subgroups—the voters who vote one way amount to one subgroup; those who vote another way amount to another subgroup; and so on. In the case before us, M. P. gives the four m's, who constitute 36 percent of the electorate, 100 percent of the elected candidates; and it leaves the four c's and the three r's unrepresented except in so far as the c's and r's agree after their differences are canceled out. If this electorate consisted of six c's and five r's, with no m's, and the candidates included at least two C's, M. P. would elect only C's. Thus Majority Preference yields, as the name implies, only the majority preference, or the preference of the minorities crudely added to make a majority; M. P. treats the electorate as a lump and fails to reflect the actual voters' preferences.

Proportional Representation, in contrast, elects the second C and the first M. Consequently, the four c's, who are 36 percent of the electorate, find themselves represented by 50 percent of the elected candidates; and the four m's, who are another 36 percent of the electorate, by the other 50 percent of the elected candidates. Only the three r's, who are nearly 30 percent of the electorate, are not directly represented. They are not so represented because only two candidates were to be elected, and those two cannot be divided to represent three subgroups. If three candidates were to be elected by the same four c's, four m's, and three r's, and

TABLE 1

TO ELECT ONE CANDIDATE

	C	M	M
c	X		
c	X		
c	X		
m		X	
m			X
m			X
m		X	
O. V.:	3	2	2

	C	M	M
c	1	3	2
c	1	2	3
c	1	3	2
m	3	1	2
m	3	2	1
m	3	2	1
m	3	1	2
M. P.:	15	14	13

TO ELECT TWO CANDIDATES

	C	C	C	C	M	M	M	M	R	R
c		X		X						
c	X		X							
c		X	X							
c	X			X						
m					X		X			
m						X	X			
m					X			X		
m						X		X		
r									X	X
r									X	X
r									X	X
O. V.:	2	2	2	2	2	2	2	2	3	3

	C	C	C	C	M	M	M	M	R	R
c	3	1	4	2	8	6	5	7	10	9
c	2	4	1	3	6	7	8	5	9	10
c	4	1	2	3	5	6	8	7	9	10
c	1	3	4	2	5	7	8	6	10	9
m	6	5	7	8	2	4	1	3	9	10
m	9	7	8	10	3	1	2	4	6	5
m	5	6	8	7	1	4	3	2	10	9
m	10	7	9	8	3	2	4	1	5	6
r	8	7	9	10	4	6	3	5	1	2
r	10	8	9	7	3	5	6	4	2	1
r	9	10	8	7	6	4	3	5	1	2
M. P.:	67	59	69	67	46	52	51	49	72	73
P. R.:		5			5					

X's represent voters' crosses. The other single letters are initials for "conservative," "moderate," and "radical," respectively. Of these, small letters stand for voters, and capital letters for candidates. Thus, in the examples, voters head rows, and candidates head columns. Each voter's row shows his choices for the respective candidates.

O. V. means Ordinary Vote; *M. P.*, Majority Preference; and *P. R.*, Proportional Representation. Each pair of examples represents the same voters' choices for the same candidates, but with election by the different methods indicated. The first pair of examples compares Ordinary Vote and Majority Preference in electing one candidate. The second pair compares O. V., M. P., and P. R. in electing two candidates. At the bottoms of the examples, the underlines mark the scores of the candidates so elected.

Each small letter could as well represent any fixed number of voters, to make the examples typify also certain large elections.

UNIVERSITY OF MICHIGAN LIBRARIES

if the r's put up a candidate, he would be elected. Thus, so far as may be when only a few candidates are to be elected, and the tellers must take the successive numerals on the ballots to represent equal steps between choices, P. R. reflects the actual voters' preferences.

The examples which we have used are rather extreme. They are not more extreme, however, than many real elections;[1] and they bring out the main differences between the several methods, suggesting wherein those differences are significant.

Which method of voting is best; or, which methods are best, and for what purposes?

Clearly, *Ordinary Vote is the best way to choose one of only two candidates.* When, however, there are more than two candidates, "the old single-shot ballot," as Hoag and Hallett called it,[1] often fails to show what the voters want. For that reason, the mathematicians Condorcet, Gergonne, Andrae, Nanson, and Hallett, the schoolmaster Hill, the barrister Hare, and others individually contributed to methods which do show what the voters want when there are more than two candidates.

Majority Preference is the way to choose only one of more than two candidates. This is because Majority Preference compares every candidate with every other candidate, all in a single election. In this way it reveals the whole electorate's first choice of all the candidates. The chosen candidate will be a conservative, a moderate, or a radical (assuming, for illustration, that these are the possible types), according to whether the electorate is preponderately or in effect conservative, moderate, or radical. Thus, in each pair of examples which has been given, M. P. elects moderate candidates because those candidates are most preferred by the electorate as a whole. The first electorate is made up of three conservative and four moderate voters, hence is preponderately moderate. The second electorate comprises four conservatives, four moderates, and three radicals. Since the radicals nearly counterbalance the conservatives, this electorate taken as a whole prefers moderate candidates to any other one type.

Majority Preference is also the way to choose more than one of more than two candidates for congeniality; congeniality of the elected candidates either with the electorate as a whole or with each other. For example, a socially homogeneous club may use Majority Preference to elect new members whom the club as a whole finds congenial. Any electorate may use M. P. to elect a mutually congenial group as a purely executive body, one which is not to develop any new policies or ideas but merely carry out policies which the electorate has fully defined. Such homogeneous clubs and executive bodies, however, seem relatively uncommon. Modern life is so complex that many social groups prefer a somewhat diversified membership; and most electorates authorize their executive bodies to make many important decisions, and at least to develop and recommend policies. Moreover, many elections are to choose members of essentially legislative bodies.

Proportional Representation is the way to choose more than one of more than two candidates for rich representation. Here the voters want to express themselves not as a single unit, which is often a fiction, but as voters whose various interests and ideas should contribute, if only as stimuli, to every decision that affects the whole electorate. Through P. R., the voters as free individuals constitute themselves into groups, whether wittingly or not, and obtain representation somewhat in proportion to the sizes of these groups. Thus P. R. expresses best the interests of the whole electorate.

Each of the methods is useful likewise for choosing between policies, plans, or other alternatives: O. V. for choosing one of two, M. P. for choosing one of more than two, and P. R. for choosing more than one of more than two alternatives.

Contrary to some current impressions, Proportional Representation is not a special instrument of and for political parties. Political parties have used P. R.; and we have heard much, including much that is erroneous, about how P. R. helps minority parties into office. A recent example was the election of two Communists to the Council in New York City in 1945. Politicians and other opponents of P. R. made the most of that election to persuade the voters to repeal P. R.; and the voters did repeal it in 1947. Apparently the voters overlooked the facts that political groups always use whatever methods of voting are available to them, and that such groups have often misused O. V. as they cannot misuse P. R.

Examples are Hitler's coup in Germany; machine rule in Kansas City as compared with popular rule in Cincinnati; and machine rule in New York City under O. V. as compared with popular rule there under P. R.[2] Indeed, before New York City adopted P. R., the Tammany organization elected, year after year, more than fifty-five of the sixty-five aldermen. One year that organization elected, somehow or other, sixty-four of the sixty-five. In contrast, under Proportional Representa-

tion, according to Hallett and Woodward, "the votes for candidates of all groups have always been faithfully reflected by P. R."[3] Naturally, any self-seeking group of politicians opposes P. R., unless the group is underrepresented under other methods.

P. R. neither creates a multitude of parties nor destroys the two-party system. In Italy, ten minority parties tried to win the general election in 1913, which was six years before Italy adopted anything like P. R. Conversely, in Great Britain, Australia, New Zealand, South Africa, and Canada, "the single-member district majority system of election" failed to save the two-party system.

P. R. does not split up the electorate into factions of any kind. When whatever groups that occur within an electorate are allowed fair representation through this system, they are likely to become less prejudiced and more cooperative. Some students of voting believe that if the several groups within the North and the South had been fairly represented in Congress in the 1860s, such representation "would at least have tended towards making a peaceful settlement possible."[1]

Instead of creating groups within the electorate, P. R. recognizes how the voters themselves divide on any given issue. Often these groups are temporary and shifting; for example, Voter Smith votes as a conservative on one issue, as a radical on another, and as a moderate on a third. Even in municipal elections, close observers report that many voters do not follow party lines but cross those lines frequently in registering choices on the ballot.[4]

Proportional Representation does not work against majority rule. On the contrary, it enables any majority of voters to elect their candidates to a majority of the offices. Without P. R., a minority is likely to rule not only through an accidental or a manipulated election process, as we have seen, but even under the guise of majority government. For example: A body of voters are, say, 51 percent conservative and 49 percent moderate. Of the total body, 26 percent, who are conservative, oppose a certain waterway project, and all the remaining conservatives and the moderates—74 percent in all—favor that project. In electing a legislature, the conservatives win all the seats. More than half the legislators represent the 26 percent of voters who oppose the waterway project. So constituted, the legislature puts through a bill that defeats the project. Thus, 26 percent of the voters thwart the remaining 74 percent. P. R., given fair opportunity for the voters, would have elected nearly as many moderates as conservatives to the legislature, and would have included among the conservatives a number who favored the project, so that, given faithful representatives, the project would have been authorized in the legislature by a vote of about three to one.

Proportional Representation protects the majority, when there is one, and minorities alike from tyranny by any group within the electorate, and it makes for enlightened majority decisions. P. R. recognizes that elections occur only because there are groups, actual or possible, many of whom have ideas which the whole electorate, or its representatives, should consider. Through P. R., the majority recognizes the minorities as part of the electorate; the majority gains from the minorities' important truths, or reveals the minorities' errors, and so strengthens the position of the whole electorate; and the majority avoids martyring any minorities and driving them underground. Moreover, through P. R., whenever a minority before an election gains enough converts to become a majority after the election, this new majority is represented by its experienced, hitherto minority, representatives; whereas this new majority under O. V. might not, and under M. P. would not, have been represented at all.

Proportional Representation, finally, does not paralyze efficiency. The historic disunity of government in France and some other countries, often ascribed to P. R., can be explained quite otherwise. Those countries did not use the method of the single transferable ballot. For many years France used majority systems, then, for a time, a block-list method which Poincaré called "a transactional, hybridous, equivocal system . . . dangerously favorable to mistakes and deception."[1] The several countries have had various party-list methods which lend themselves to political manipulation and hamper the voter's freedom as what we are calling P. R. does not. Even the party-list methods in vogue in those countries, however, have brought many advantages over the cruder methods of voting which they displaced.[1]

A legislative body, a board, or a committee which is elected by P. R. can often work out plans which combine the insights and interests of the various groups within the electorate. Hoag and Hallett pointed out that "the P. R. councils of Cincinnati and Sacramento were able to choose the first city managers of those cities by unanimous vote."[1] Whenever the elected representatives of different points of view cannot agree on a plan, they can put forward their several plans and choose one of these, or have the electorate choose one, by M. P.

UNIVERSITY OF MICHIGAN LIBRARIES

Of course, none of the methods works perfectly. Each method, to work well, requires honest and sufficiently able candidates, voters, and tellers. Each method can be hampered, for example, by "strategic" voting—voting not according to one's preferences for candidates but according to one's estimates of candidates' chances for election; though this distortion is impossible in O. V. for one of only two alternatives, and is unlikely in either M. P. or P. R. Moreover, the numerical intervals between the voter's choices on his ballot often fail to reflect the real intervals between those choices in his mind. Still more accurate methods than the three that have been described are possible under some circumstances. To apply each of the three methods, however, to the special conditions mentioned—O. V. for election of one of only two alternatives, M. P. for election of one of more than two, and P. R. for election of more than one—seems to be the most workable way to improve elections themselves; and to so improve elections is to encourage suitable nominations, voting, and consequent action.[1,5]

For elections by more than a few hundred voters, of the three methods, only Majority Preference is seriously unwieldly. It is so greatly needed, however, that someone ought to invent a voting machine that would make this method convenient for large elections.

For smaller elections, each of the methods is readily workable. True, when the ballots are few, the usual way of counting the ballots for P. R. introduces chance elements which may give somewhat different results on a recount of the same ballots. An exact way to count the ballots is therefore essential. Such a way is available; a way whereby every accurate count or recount of the ballots comes out the same, even for an election to fill two offices from four candidates by four voters who may themselves be the candidates.[6]

It follows that any voting body, an association, a faculty, a department, a board, or a committee, may well adopt the rule that "elections of one of two alternatives shall be by ordinary vote; one of more than two alternatives, by majority preference; and more than one of more than two alternatives, by proportional representation with the single transferable ballot."

REFERENCES

1. Hoag, C. G., and Hallett, G. H., Jr. *Proportional Representation*, 1926, 13–33, 2, 30, 435, 60ff., 140, 100–10. New York: Macmillan.
2. Hallett, G. H., Jr. *Proportional Representation: The Key to Democracy*, 1940, 7, 2, 18–20. New York: National Municipal League.
3. Hallett, G. H., Jr., and Woodward, W. R. *Nat. Municipal Rev.*, 1947, **36**, 287.
4. Articles by Hallett and others in *Nat. Municipal Rev., passim.*
5. Hallett, G. H., Jr., and Woodward, W. R. *Forum*, 1947, **107**, 97–104, 392–8.
6. For large elections by M. P., cf. Hoag and Hallett, *op. cit.*, 496–505; by P. R., inquire of National Municipal League, 299 Broadway, New York City.
   For small elections, cf. the foregoing sources or, for what some find more convenient and exact methods, W. S. Taylor, *Rules for . . . Majority Preference*, and *Rules for . . . Proportional Representation* (both mimeographed), 15 Pierce Hall, Northampton, Mass.

SCIENCE ON THE MARCH

TROPICAL MEDICINE—A PARTIAL REPORT

AS A nation, we as North Americans have manifested little interest in tropical diseases, although they cause more than one half the illnesses and deaths suffered by mankind. With the exception of a small group of doctors and technical men, this attitude is true also of the medical and allied professions.

During the first world war there was a flurry of interest, which soon subsided. During the second world war, however, when it was necessary to send large contingents of our soldiers into tropical regions, the nation as a whole began for the first time to realize in a small way what was meant by tropical diseases and to understand what a great influence they have on the world's economy and, indirectly, upon our life here. During the war years this interest stimulated intense study of these diseases, and it is gratifying to know that that interest continues.

Notable advances that are of extreme interest both to the professional man and to the general reader have been made during the past several years. To my mind, the most absorbing story in this field is the story of the development of chemotherapy in the treatment of Hansen's disease (leprosy). Only a small percentage of the people in the United States are familiar with the ravages of this chronic and fatal affliction, but everyone has some idea (usually mistaken) about its results and dangers.

The story of the new life interest and hope that has been brought to inmates of leprosaria with the advent of sulfone treatment has never been written. These sulfones are chemically related to sulfanilamide and sulfathiazole. The initial study was begun by the late Dr. Faget at the National Leprosarium, Carville, Louisiana. Early in 1940, he began using Promin in treatment of a group of patients in that hospital. Three years later, after careful study and observation in order to be sure that he did not make a falsely encouraging report, he published a preliminary report of his results. Since then different sulfa drugs have been tried out in various hospitals throughout the world, with uniformly good results. At present it is undecided whether sulfone therapy results in a permanent cure. The question is of technical interest only, however, since so far as the patient is concerned the progress of the disease is stopped and there is the possibility that he will become bacteriologically negative, receive a discharge from the institution, and return to his family and friends. I am glad to report that the International Leprosy Congress held last April in Havana voted the sulfones to be the treatment of choice for Hansen's disease.

The world in general little realizes what this means to a patient with Hansen's disease, and the written word is powerless to convey even a remote idea of the hope it brings to these formerly condemned individuals. Anyone who visited the different leprosy hospitals before 1940 and saw the hopeless attitude of the patients, the masklike, apathetic faces, the dull eyes from which the light of hope, ambition, and desire for life had been extinguished; who visited countries where sufferers were chased out of villages and forced to live like animals in the jungle; or who saw them isolated in deep pits from which they could not climb, but were fed by meager rations lowered at the end of a rope (knowing that when merciful death at last arrived the hole would merely be filled up)—only such a person can have some idea of the humanitarian side of the story back of such a discovery.

It has been my privilege to visit several leprosaria where the sulfone drugs had been used during the previous twelve months—the change was astounding! The characteristic ragged ulcers were healing, the nodular, disfigured faces were returning to normal, and many of the beginning cases of blindness could see again. Many patients who had been complete invalids were going about assisting in the work of the hospital colony. It is quite true that because the drugs are of recent origin there remain numerous old cases of Hansen's disease with mutilated faces and extremities, but the entire mental atmosphere is changed. You now meet patients who know they have an opportunity to get well, who realize that they have a chance to discard the fictitious names—adopted in order to protect their families—and to return to their homes as cured, or arrested, cases.

Those who visit one of the larger colonies leave with the memory of the unfortunate children, sometimes as many as five hundred, who are confined in a single colony because of Hansen's disease. These patients, some of them only five years

UNIVERSITY OF MICHIGAN LIBRARIES

old, do not as yet realize what the future holds in store for them—but the visitor can see on all sides the misshapen, the blind, and the disfigured derelicts into which these children will develop should they be denied sulfone therapy.

To illustrate "man's inhumanity to man," one has to seek no further than the treatment of these unfortunate people. The fear and dread that the public has manifested are out of all proportion to the danger involved. I believe any doctor familiar with the disease will agree that a person with active tuberculosis is a far greater danger to the community than the patient infected with leprosy. In fact, the mode of transmission of this disease has not yet been discovered or demonstrated, although well over a hundred human inoculations have been performed without a single positive result.

The most recent advance in the field of tropical medicine was announced in two reports given at the International Congress of Tropical Medicine in Washington, D. C., May 11 of this year. One described the work of Smadel and associates who studied the use of Chloromycetin in treatment of scrub typhus; the other report was the study done by Payne and Knaudt, who treated epidemic typhus in Bolivia. These two dreaded diseases are caused by minute organisms belonging to the *Rickettsia*.

The rickettsial diseases, which often range far north, have continuously defied medical science to offer anything in the way of a cure or remedy. Chloromycetin, a new antibiotic, was tried with spectacular success in these two diseases. Epidemic typhus, which is transmitted by the human louse, has had a great influence on the history of the world—perhaps second only to malaria. It is a disease that has often accompanied armies, devastating and destroying them before they reached or accomplished their objectives. And it remains after the army has passed, to add its horrors to the starvation and wreckage of war. Where typhus fever is endemic the fear of the disease testifies to its seriousness. The mortality rate, varying from 20 to 50 percent, bears witness to its destructiveness.

Scrub typhus is perhaps the most serious disease of the East Indies. During the recent war with Japan, this disease at times caused more casualties than the enemy. It was a serious threat to successful military operations. Mortality from scrub typhus is high and has rendered considerable territory uninhabitable throughout the Orient. It is transmitted by a small insect, a larval mite, which has the musical name *Trombicula* and is so small that sixty-four of them standing bumper to bumper would be only one inch long.

The greatest scourge of mankind, however, has been, and continues to be, malaria, from which approximately 10 percent of the world's population suffers each year. Formerly the only remedy available was quinine. The world never produced more than 600 tons of this drug per year; this was used not only for the treatment of malaria but also in the preparation of tonics and even of cosmetics. Yet to treat all cases of malaria the world would each year need 4,000 tons of quinine. During the first world war, Quinacrine (commonly known as Atabrine), the first substitute for quinine, was discovered. It has therapeutic advantages over quinine but also has the disadvantage of imparting a yellow color to the skin. Although this discoloration does not as a rule have any effect on the well-being of the patient, nevertheless it is the basis for many objections to taking the drug. During World War II, because of the immense importance of malaria to military activity in tropical regions, several cooperative studies of the disease itself were made in different countries and intensive search was made for new and better curative drugs.

An outstanding study, organized under the direction of Sir Hamilton Fairley at Cairns, Australia, was a monumental work in all phases of malaria research. In North America the malaria section of the Office of Scientific Research and Development carried on a similar study. These studies resulted in the discovery of several new compounds, some of which have already been added to the therapeutic agents for treatment of this disease.

These new synthetic drugs are more efficient than either quinine or Atabrine. In addition, they do not discolor the patient's skin. With these new drugs for treatment, and the discovery that DDT is highly effective against certain species of the *Anopheles* mosquito, which transmits malaria, it is now possible to carry on commercial and agricultural enterprises in many formerly deserted regions.

The use of penicillin in conjunction with arsenicals has reduced to a few days the duration of illness due to yaws. Yaws affects, as a rule, only the poorer populations and is one of the most loathsome diseases encountered in tropical regions. Kala azar, found in Mediterranean and Oriental countries, now can also be successfully combated with synthetic compounds.

African sleeping sickness, or trypanosomiasis, has been the subject of a great deal of work and

study. This disease, which has depopulated large areas of the African interior, attacks domestic and other animals, rendering stock raising impossible in regions where the disease is endemic.

Progress has been made in the therapeutic field as well as in the field of prevention. Study of the habits of the vector, the tsetse fly, has enabled control measures to be put in practice which have greatly reduced the fly population. In the field of chemotherapy, a group of drugs has been found which forms a fairly successful combination. Bayer 205 is used as a prophylactic, amodines are used in treating early cases, and Tryparsamide in treating late or neurological cases. Thus the problem of one of the most contrary and baffling diseases of the tropics has been at least partially solved.

One disease in which no recent therapeutic advance has been made is amebiasis. The protozoa responsible for this disease, which can cause serious illness and sometimes death, may live for years in the human body without causing trouble for the host, as is evidenced by the fact that 10,000,000 people in the United States harbor these parasites, yet few suffer serious results. Although recently much time, effort, and money have been expended in the search for more effective treatments, nothing new has been developed that constitutes an improvement on classical, established treatment.

Recent studies have taught us a great deal, however, about the life and habits of *Amoeba histolytica,* and such information is necessary for a properly organized attack against such a parasite. In this way progress is made in the war against disease. Any new discovery is the result of a long series of studies by numerous primary researchers who deserve, but seldom receive, the plaudits of an adoring, appreciative public.

The student entering the field of tropical medicine need not fear that nothing remains for him to discover. Tropical medicine contains many mysteries—numerous diseases for which no cure is available and many others whose etiology is unknown. Tropical ulcer, which incapacitates so many in the Amazon Valley and is especially severe along the Rio Abuna, occurs in varying severity throughout the tropical world. Its cause is in doubt, and no effective satisfactory treatment is available. Chagas' disease, or South American trypanosomiasis, although its cause is known, still remains the great therapeutic mystery of this hemisphere. No drugs so far have given any indication of activity. The cone-nosed bugs that transmit the disease are not easily controlled even with DDT. The student will be intrigued by the highly fatal Oroya fever, with its romantic history; by filariasis, with its startling sequela of elephantiasis; and Espundia, with its horribly disfigured victims, will stimulate his search for an effective cure.

These are only a few of the problems that remain in the heavily populated tropical countries.

EUGENE H. PAYNE

Parke, Davis & Company
Detroit, Michigan

PETROLEUM AND NATURAL GAS AS SOURCE MATERIALS FOR CHEMICALS

PETROCHEMICALS were a negligible factor in our economy in 1925. Annual production was less than 150,000 pounds. Today they comprise a mammoth industry which produces over 3,500,000,000 pounds of raw materials for synthetic organic chemicals yearly. This development also represents a marked increase in the percentage of aliphatic chemicals produced. In the 20-year period ending in 1945, aliphatic production rose from about one third of all organic chemicals to two thirds of the total. Petroleum and natural gas, however, are becoming increasingly important in supplying aromatics because present requirements are not being met by carbonization of coal. Although petroleum and natural gas are in extremely high demand as fuels, the quantities needed to expand chemical production are very small percentages of their total. The present high tonnages of petrochemicals require less than 1 percent of the total petroleum production.

Even abroad, where petroleum is for the most part imported, construction is under way for increasing petrochemical production. In England, the first installation of a high-temperature cracking process is now expected to have a capacity of about 150,000,000 pounds annually. Other plants for petrochemicals in England will provide for an annual output of 100,000,000 pounds of detergents, 60,000,000 of insecticides, 60,000,000 of ethylene, and 48,000,000 of olefins, aromatics, and oxidation products. A new chemical industry based on petroleum is also being established in Holland. Plants now near completion will produce 4,000,000 pounds of plastics, 50,000,000 of detergents,

UNIVERSITY OF MICHIGAN LIBRARIES

and over 100,000,000 pounds of nitrates. Other countries are also considering installation of petrochemical plants.

The exponential growth of the petrochemical industry has been brought about largely through chemical research. It is interesting to note that more than 5,400 finished products are now made from petroleum. Many of the raw materials are by-products of the streamlined refinery processes used in motor fuel production. Catalytic and thermal cracking, hydrogenation, dehydrogenation, polymerization, isomerization, alkylation, and superfractionation make possible the production of individual hydrocarbons of high purity. The oxidation of hydrocarbons to carbon monoxide and hydrogen is one of the most versatile used in the petrochemical industry. Among its products are many types of oxygenated products and higher molecular weight hydrocarbons. The American Fischer-Tropsch process, which is being developed primarily for the manufacture of motor fuel from natural gas, will also produce large quantities of alcohols, aldehydes, acids, and ketone. Production of these compounds from the new plants now projected has been estimated as high as 300,000,000 pounds per year. Improved processes for direct oxidation of natural and refinery gases and others are continuously increasing the number and amounts of products. A process has been worked out for the thermal conversion of natural gas or components thereof into acetylene, and a huge plant is being installed to carry out the process. Acetylene can also now be produced from natural gas by a "glow discharge electrical process." The first commercial plant, located at Houston, Texas, will be completed this year. A high-temperature cracking process perfected in England, in which olefin gases and aromatics are produced catalytically from petroleum distillate, is a new development. Even before the first plant is completed, research has brought about improvements which will result in a 50 percent increase in capacity over the initial estimates.

The hydrocarbon raw materials derived from petroleum are principally ethylene, propane, propene, butanes, butenes, butadiene, pentanes, pentenes, pentadienes, toluene, and xylenes. Large amounts of the cresylic and naphthenic acids are also present in petroleum. The cresylic acids have a variety of commercial applications in resins, adhesives, surface coatings, gum inhibitors, flotation agents, disinfectants, metal cleaners, solvents, and insecticides. The naphthenic acids are used in fungicides, high-pressure lubricants, and rust preventatives. Their copper and zinc derivatives are used as paint driers and for treating textiles, in which they prevent deterioration from mildew and mold.

For the past few years, synthetic rubber has been the largest single consumer of hydrocarbon chemicals. In 1946, over 1,000,000,000 pounds, or about one third of the total petrochemical production, was butadiene, all of which was allocated to the synthetic rubber industry. Large amounts of styrene, isobutylene, and other hydrocarbons are also used for this purpose. The availability of substantial quantities of natural rubber probably will not result in a decrease of petrochemical production, for these compounds are in demand for other products, such as plastics and resins. Second, new rubbers are being invented that are far superior to the natural for many purposes. Butyl rubber is a good example. As a material for inner tubes, Butyl is far better than natural rubber because diffusion of air is about one tenth of that through natural rubber, and it has unusually good age and tear resistance as well. Research also indicates that Butyl rubber is superior material for tractor tires, where deterioration from oxidation and light are important factors. A recently developed synthetic rubber has properties that make it superior for tires over previous types. In numerous road tests, tires of this rubber have outworn the natural.

The agricultural industry is being revolutionized by chemicals from petroleum and natural gas. The oil industry is engaged in extensive research on a wide range of products from pesticides to preservatives. An outstanding development for the extermination of nematodes is the soil fumigant, D-D (dichloropropane and dichloropropene), which has resulted in increasing production of some crops as much as 200 percent. Weed killers of the synthetic hormone type are being made from by-products of catalytic cracking. Indications are that petroleum will play an increasing part in supplying insecticides as the necessity for synthesizing aromatics from petroleum becomes greater. DDT and Gammexane, for example, will be made from petroleum. The industry is now doing a $100,000,000-a-year business in insecticides and fungicides. The newly developed weed killers are expected to provide an even greater volume of business for petroleum chemicals. Other petroleum products which conserve food supplies are ethylene and higher gaseous olefins, used to color citrus fruits and hasten the maturing of peaches, apples, cherries, plums, and other fruits.

Another regulating synthetic chemical is "endrop," or alpha-naphthalene acetic acid, which prevents apples from dropping prematurely.

The detergent field is one of the principal markets for petrochemicals. It is in a state of rapid growth and is expected to reach 1,000,000,000 pounds annual output in a few years. Total production of synthetic detergents in 1947 was about 600,000,000 pounds, which contained about 240,000,000 pounds of surface active agents. The general types derived from petroleum, which include the alkyl aryl sulfonates, alkyl sulfonates, and nitrogen-containing aliphatics, constitute a high percentage of the total produced. For many purposes, the synthetic detergents are superior to soap. They are more effective in hard water, can be used in acid solutions, and do not leave curd, which causes such annoyances as "rings in the bathtub." Two factors that have inhibited more widespread replacement of soap have been the inability to produce them in bar form and the comparative ineffectiveness of detergents on heavily soiled cotton goods. Recently, the bar formation problem has been solved for some detergents, and they are now being marketed in that form. Research indicates that the cotton problem can be solved by adding small amounts of phosphates, carbonates, borax, or carboxymethylcellulose. A group of surface active agents, of which quaternary ammonium compounds are the most important, possess excellent germicidal properties. Synthetic detergents also have a variety of applications in textile finishing. Other interesting uses include settling coal mine dust and washing dirt from city streets.

The plastics and resin industries comprise one of the largest markets for petrochemicals. Many types such as polystyrene, vinyl resins, and polyethylene are produced either wholly or in part from petroleum. The 1947 consumption of polystyrene, which requires large amounts of ethylene, was 92,000,000 pounds. It is expected that capacity will be expanded to 200,000,000 pounds by the end of this year. Petroleum also supplied a high percentage of the material for 180,000,000 pounds of vinyl resins produced in 1947, and demand is not yet satisfied in this field. This year's production of polyethylene is anticipated to be 50,000,000 pounds. Petroleum is also becoming a growing factor in phenolic resins and plastics, of which over 200,000,000 pounds were produced in 1947. The demand for phenol for all purposes is much higher than production at present. Indications are that the plastics and resin field will require increasing amounts of chemicals for a number of years. Through research, new adaptations and new products are continually being invented. These products, which range from textile finishes and fibers to radio cabinets, make possible such innovations as water-repellent felt for hats, glass that sheds ice and moisture, and unbreakable phonograph records.

Petroleum also supplies raw materials for a host of other synthetic chemical intermediates and solvents. The protective-coatings industry, for example, utilizes a number of these products for various purposes. Such solvents as acetone, methyl ethyl ketone, isopropyl alcohol, and methyl isobutyl carbinol are derived from petroleum gases—ethylene, propylene, and butylene. Polyhydric alcohols have been developed which improve the drying and durability properties of varnishes. Synthetic glycerin and its by-products, allyl chloride, allyl alcohol, epichlorohydrin, also have numerous applications in the coatings industry. The initial plant has a capacity of 35,000,000 pounds of glycerin per year. Another development of major importance is the production of phthalic anhydride from *ortho*-xylene. The first commercial installation is producing at the rate of 8,000,000 pounds a year. This process will undoubtedly be expanded because naphthalene, formerly the sole basic chemical for phthalic anhydride, is not meeting present requirements.

From a tonnage standpoint, synthetic methanol is one of the most important products made from petroleum and natural gas. In 1947, total production was 545,000,000 pounds. A great increase in 1948 is indicated by the fact that almost 73,000,000 pounds were produced in March alone, compared to a maximum monthly production of about 48,000,000 pounds in 1947. The trend toward utilizing petroleum and natural gas as raw materials is illustrated by the shift in sources for synthetic methanol. In 1946, 71 percent was based on coke or coal, whereas, in 1948, 77 percent is expected to come from natural gas.

An important consequence of the expanded methanol production is the increasing quantities of formaldehyde available. Maximum monthly production in 1947 was 46,000,000 pounds; in March 1948 it reached over 60,000,000.

Although complete U. S. data on synthetic ethyl alcohol are not available, it has been estimated that the annual production is about 200,000,000 gallons and that over 50 percent of this volume comes from petroleum and natural gas. It is made largely through reaction of ethylene with sulfuric acid followed by hydrolysis. Commercial installations are also being made to utilize a catalytic process for hydration of ethylene to ethyl alcohol in

UNIVERSITY OF MICHIGAN LIBRARIES

…t with an annual capacity of 18,000,000 …ons.

…ther high tonnage chemicals are manufactured …er wholly or in part from petroleum and natu… …gas. Among those in over 100,000,000 pounds …ual production are acetaldehyde, acetic acid, …ic anhydride, acetone, butyl alcohols, carbon …achloride, ethylene glycol, and isopropyl al…ol. It is estimated that of the total domestic …duction, 35 percent of the normal butyl alcohol, …ercent of the acetone, and all the isopropyl …hol come from petroleum and natural gas.

…etroleum and natural gas also furnish ma…als for certain inorganic chemicals. Large …ntities of sulfur compounds are present in …y crude oils and natural gas. As an example, …00 tons of elemental sulfur were recovered …years from an Arkansas natural gas contain…6-8 percent hydrogen sulfide. In a number of …eries, hydrogen sulfide from processing crude …idized to sulfur dioxide for conversion to sul… acid. Hydrogen is obtained from the high…erature cracking of natural or refinery gas. … of particular importance because it is used to …ufacture ammonia and ammonium nitrate, …h are so important in meeting fertilizer de…ds. One particular plant produces sufficient …ities to meet 11 percent of the United States …irements.

…he outlook for petrochemicals is extremely … despite the competitive demands on petro… and natural gas for fuel purposes. The pic… in relation to other raw material sources for …icals is so clouded with acute shortages and … costs that the necessity for increasing use of …oleum and natural gas is evident. As a supplier …iphatic chemicals, fermentation of agricultural …ucts is the largest competitor. Prices for …, however, are such that it is more valuable …od for humans and livestock than for fermen… purposes. There is also strong competition for all fats and vegetable oils as food products. The cost of aliphatics from coal, starting with acetylene, is excessive because of the price of coal. Wood distillation, once the second largest source of aliphatics, cannot be expected to furnish the large-scale demands of today from our dwindling forests.

In the aromatic field, petroleum is becoming continually more important. Coal tar, the principal source of aromatics, is limited by the extent of by-product coke-oven operations in the manufacture of steel. Demands have already greatly exceeded the capacity of this industry. Many of the new petrochemical plants now under construction will make large quantities of aromatics. For example, one projected plant on the Gulf Coast is expected to produce as high as 200,000 pounds of aromatics per day from natural gas.

Plant construction by major oil and chemical companies now under way is in itself indicative of the future of the petrochemical industry. Developments in the Gulf and West Coast areas are particularly significant. On the Gulf Coast alone, about $400,000,000 is being invested in chemical plants. Announcements of further construction are continually forthcoming, and this estimate is conservative. On the West Coast, one oil company plans expansion to manufacture 300,000,000 pounds of detergents annually, about 30 percent of the future market. Another oil company will expand its Western operations to increase ammonia production by 50 percent and to manufacture new organic chemicals. Other Western companies are also increasing petrochemical output. These are but a few examples of the skyrocketing growth of the industry. The foundations that are being laid for the petrochemical industry attest its sound position in our economy.

GUSTAV EGLOFF
Universal Oil Products Company
Chicago, Illinois

BOOK REVIEWS

BASIC TOOL

The Genius of Industrial Research. D. H. Killeffer. vii + 263 pp. $4.50. Reinhold. New York.

PERHAPS it is the publisher's fault, perhaps it is Mr. Killeffer's, but the fact is that the latter should not appear as "author" of this book. Call him editor, collator, or compiler, but certainly not author, for more than 60 percent of the book consists of direct quotations from the writings of others.

Just why Killeffer does not appear in his true guise is hard to understand. Listed as editor, he would receive ample credit, for he has done an excellent job of collecting and assembling material to illustrate and prove the points he wishes to make. At least ten Perkin Medalists and more than a score of other scientific notables relate their experiences and contribute advice.

The book is one of several which might well be required reading in a course that should be given every third- or fourth-year science student.

A thorough understanding of the methods of research is one of the basic tools, the hammer or saw, that must be in the kit of the young research man. Without it he can build no permanent structure but only odd bits of ill-assorted bric-a-brac. Yet personal experience and observation of some hundreds of chemists lead the reviewer to believe that a majority escape college with little real conception of the methods they must use if they are to do creative work. Killeffer puts it bluntly: " . . . methods of research and development as practiced in our industries are not taught in our colleges and universities. The average graduate or postgraduate going into research has to learn, if he ever does, by hard experience the simple principles of his art."

The fault is that of uninspired, or ignorant, instructors, and a book such as Killeffer has edited could do much to direct the undergraduate gaze toward unsuspected vistas. Surely no better examples of the clear, logical reasoning that leads to great discoveries can be found than in some of the great classics of chemical research. Among those Killeffer has included are the story of Carothers and synthetic rubber, Langmuir and the gas-filled lamp, Midgely and lead tetraethyl, Sabatier and catalytic hydrogenation, Hall and aluminum, Hyatt and celluloid, Dow and bromine, Cottrell and the dust precipitator, Williams and thiamin.

The book contains a number of good bits of advice for the research man. Typical are:

> Although it is difficult to persuade research men to admit such a thing, it is possible to provide too much ease and too great convenience for the researcher. The process of incubation and relaxation may become so pleasant that the time of the researchers may be so wholeheartedly devoted to this phase of the operation . . . nothing at all is accomplished.
>
> The process of research reduces itself to a succe of statements of the problem ever more precisely foc upon the solution as information accumulated bit b shrinks the unknown area around the objective. . . vital [to the researcher's] own efficiency that he dra frequent reports to himself, if not to others.
>
> Far too often the research man abandons his bra the door of the pilot plant, and because he understand it is the place to make mistakes, proceeds to make reasonable error and some unreasonable ones.
>
> . . . most research workers [are reluctant] to fac fact that they often do not know what they mean intend to say when they come to write a report. . . leads to the conditions most often complained o scientific writing: that meaning is hidden or lost en among useless words.

And any research director may profitably rea chapter dealing with the proper way to state a p lem. As Editor Killeffer says: "Every research oratory is cluttered with problems left unsolved. sumably these are beyond or outside the abiliti skills of the available workers, but much more l they have not been stated in practicable terms."

RALPH F. W

Research Department
Columbia Chemical Division
Pittsburgh Plate Glass Company
Barberton, Ohio

THE STORY OF THE NACA

Frontiers of Flight. George W. Gray. ix + 30 pp. Illus. $6.00. Knopf. New York.

THE characteristic ability of the author t scribe highly technical matters in a lan intelligible to the scientifically untrained layma admittedly the secret of the readability of this b The ease with which the reader comprehend many functions of the research establishments o National Advisory Committee for Aeronautics s to impress upon him the numerous and varied fa that must be considered in designing present-da craft.

This story of the NACA is the story of the s tific development of aviation in the United S Fostered by Dr. Charles D. Walcott, then secr of the Smithsonian Institution, the idea of a go ment-supported aeronautical research organiz was unpretentiously realized in 1915 when Con passed the Navy Appropriations Bill. Attached rider to this bill was an act which empowered President of the United States to appoint a comm whose function should be "to supervise and direc scientific study of the problems of flight, with a to their practical solution."

UNIVERSITY OF MICHIGAN LIBRARIES

The Act specifies that in appointing the Committee, the President should give representation to certain groups naturally interested in aeronautics. Thus, the Committee is made up of two members from the Army; two from the Navy; two from the aeronautical activities of the Department of Commerce; one each from the Smithsonian Institution, the National Bureau of Standards, and the Weather Bureau; and six additional members from the ranks of "those acquainted with the needs of aeronautical science, either civil or military or skilled in aeronautical engineering or its allied sciences."

This Committee is the controlling body of the government's aeronautical research establishment, which today represents an investment of $85,000,000, located in three laboratory centers: Langley Field, Virgina; Moffett Field, California; and Cleveland, Ohio. Members of the Committee serve without pay and have been of such high character and distinction as to render the Committee completely free from political influence.

Employing a staff of scientists and research engineers (wartime peak—6,804 people), the NACA has undertaken and solved countless problems relating to the fundamentals of aircraft design and performance. This work is carried on not in competition with private industry, but in support of it. As rapidly as findings are made and confirmed they are made available to the Army, the Navy, the Civil Aeronautics Authority, the aircraft industry, and the engineering profession. Scientific study and experiments carried on by the NACA have resulted in improved designs for seaplane hulls, engine cowlings, high-lift wings, high-speed wings, propellors, control surfaces, wing panels, engines, superchargers, and many other items essential to more efficient aircraft.

Mr. Gray has set down an illuminating narrative of events concerning the establishment and the subsequent growth of the NACA; but the larger portion of the book is rightfully devoted to lucid descriptions of some of the more important findings of this capable research organization, which has contributed more to the success of American aviation than any other one group.

The author has brilliantly portrayed the pioneering role of the NACA in exploring and conquering the manifold "frontiers of flight."

J. H. Bell, Captain, USMCR
U. S. Naval Photographic Center
Naval Air Station, Anacostia, D. C.

REAPPRAISAL OF DARWINISM

Charles Darwin. Jean Rostand. 237 pp. Illus. 220 fr. Librairie Gallimard. Paris, France.

A SINGLE volume on Darwin which includes not only a short biography and a character sketch, but discussions of his major works and theories, of his predecessors and successors, his philosophy of biology and the present-day status of Darwinism—a single volume which includes all these subjects (each of which have occupied volumes), yet without skimping them with brief one-paragraph summaries, is a marvel of concise writing. The French, as has often been observed, seem to be best at this sort of thing, although it can be done well in any language if the author is sufficiently master of his subject. As the author of several works on the history of biological theories, the lives of insects, and of formal texts, Professor Rostand has the command of his subject that makes the concise treatment possible. There is at present no similar book in print in English, although there should be, for Darwin and his work deserve reappraisal every few years in the light of our changing theories and increasing knowledge. Such a book can serve two or three college generations before it must be written again. In the meanwhile, there could be no better volume than this for the student in biology who is also trying to learn French as one of his tools.

Joel W. Hedgpeth
Institute of Marine Science
The University of Texas
Port Aransas

MORPHOGENESIS AND ADAPTATION

Organic Form and Related Biological Problems. S. J. Holmes. vii + 169 pp. $5.00. Univ. of California Press. Berkeley and Los Angeles.

IT HAS been said that scientific research is 99 percent perspiration and only 1 percent inspiration. Laboratory drudgery leaves little time for thought. Trees and even bushes obstruct the view of the forest. It is therefore gratifying that some scientists are willing and able to undertake the work of thinking for themselves as well as for their less fortunately situated colleagues. Professor S. J. Holmes, in his early eighties, is one of the deans of American biology. A theoretical book from his pen definitely merits respectful consideration.

The theme of the book is stated in the first sentence: "The ability of living creatures to effect their own construction and to keep it in a tolerable state of repair presents a problem which has long baffled the efforts of the naturalistic explainer." This is the biological riddle of the Sphinx—the riddle of evident purposefulness in a machine compounded of evidently purposeless physicochemical processes. Professor Holmes does not offer any new or complete solution of the riddle, but he does give a critical discussion and evaluation of the proposed ways toward such a solution.

The riddle has at least two parts. First, there is the morphogenesis that results in the origin of a smoothly functioning and often very complex system from an egg or a rudiment which appears to have a much simpler and less active organization. Second, there is the direct adaptation, which results in the body's

altering its structure or activity to meet changes in its environment. Both parts can be considered in two aspects.

One may describe the phenomena of morphogenesis and of direct adaptation in terms of physiological, and ultimately physicochemical, reactions. The origin of these phenomena may also be analyzed and explained causally in terms of the historical development of a given organism and of the living world as a whole. Professor Holmes' book is largely devoted to the first of these two aspects, although he is fully aware of the paramount importance of the second. To use his own words, "In response to the question how the parts of the organisms came to have this fortunate endowment, the mechanist has no other consistent recourse—although many have tried to find one—than to fall back upon the interpretation of adaptation supplied by Darwin." To pursue the subject further he would have to enter the field of modern population genetics and evolution. This he refuses to do, although repeated references to Darwinism are made throughout the book. Quite consistently, the central idea of the book is that "Each step in morphogenesis may be regarded as a matter primarily of functional adjustment, a part of the general process of homeostasis, or balancing, as much as the various processes subsumed under this name in the adult body." But homeostasis is evidently not comprehensible except as a product of the kind of natural selection that Schmalhausen has recently called "stabilizing selection."

In a very thoughtful and carefully balanced chapter on Autocatalytic Enzymes and the Origin of Life, Professor Holmes develops one of probably the most fruitful ideas current in modern biology, namely, that phenomena of self-reproduction or self-synthesis are among the basic manifestations of life. But he stops short of the conclusion that the origin of self-reproduction constitutes the origin of life in the world. Natural selection, and consequently adaptive and progressive evolution, are implied in self-reproduction. He accepts, however, the dubious theory of Oparin, according to whom complex organic substances, perhaps even proteins, appeared on earth spontaneously before the dawn of life. The chapter on Cancer as a Biological Problem is somewhat isolated in the book, but it gives an interesting discussion of recent hypotheses which attempt to explain the nature and origin of malignant growths. Two of these hypotheses are favored: that of somatic mutation and that of "evocation" resembling cellular differentiation in the normal development. Professor Holmes concludes that "The somatic gene mutation theory and the evocation theory may not be so wide apart as they may appear. It is quite possible [he believes] to account for the facts favoring both of them by a somatic mutation theory which attributes the origin of malignance to changes not in genes but in gene derivatives." Provided that the "gene derivatives" are self-reproducing bodies, their change which is perpetuated in their division products is a mutation by definition. To a geneticist, the somatic mutation theory does not necessarily mean that a gene borne in a chromosome must be changed. This theory furnishes, then, the most hopeful path of approach to the problem of the origin of neoplastic growths.

TH. DOBZHANSKY

Department of Zoology
Columbia University

THE MOVING FINGER WRITES

Our Plundered Planet. Fairfield Osborn. xiv + 217 pp. $2.50. Little, Brown. Boston.

MORE significant than a political and military history of the world is this moving account of the impoverishment of the earth's biological balance. Man's history of self-destruction as written in the deserts of Asia, the Mediterranean lands, and Africa, in Europe, and even in Russia and Australia is being rewritten at an accelerated pace in the continents of the New World. The pattern is unalterably the same—misuse of the land. Despite the strengths of their armies, the acumen of their mechanical and industrial geniuses—in fact, often because of them—nation after nation has fallen into decay through ignorance and lack of cooperation with the natural forces of the planet.

Fairfield Osborn's question: "Are we to continue on the same dusty perilous road once traveled to its dead end by other mighty and splendid nations, or, in our wisdom, are we going to choose the only route that does not lead to the disaster that has befallen so many other peoples of the earth?" remains unanswered. To understand the question in its full and awful meaning, one must study this most disturbing history. To answer the question hopefully, statesmen and scientists together must devise positive action to stem the flow of lifeblood—soil and water—from a rapidly dying and highly complex biological entity, the planet Earth.

In writing the history, in pointing the finger of effect at the causes of the world's No. 1 problem, the president of the New York Zoological Society has performed a task equal in importance to that of the prophets of old.

A. G. HALL

The American Forestry Association
Washington, D. C.

UNIVERSITY OF MICHIGAN LIBRARIES

TECHNOLOGICAL NOTES

Distinctive subject. Scientists who pursue photography for the love of it or for its use in their work are generally "up" on what is new in equipment and techniques. Manufacturers' releases on cameras, accessories, and methods are relatively few and highly technical.

On the other hand, nearly every release, regardless of subject, is an example of the use of photography in industry. Pictures supplement the written word or even replace it. Technical description by means of photographs is comparable to the felicitous use of language to explain what and how, and to some extent why. Improvement in this form of expression is evidenced in concentration on the technical objects themselves rather than making them merely incidental to the beauty of the model holding what is to be described.

The same sort of serious economy in presentation appears in motion pictures of industrial subjects. A subject worth presenting is worth immediate and concentrated attention instead of being dragged into a story by the heels.

History. Tracing the beginning back to Edison's Kinetoscope of 1889, the Royal Bank of Canada has a letter on movies in education and industry. The letter differs from the comment above: "The informational film contemplated by business and industry must be sharpened by human emotional interest; they will find information-giving a dangerous business if the emotion-kindling process is not incorporated." Yes, but sheer interest in a fascinating subject is a sign of kindled emotion.

What Radio Corporation of America calls "television history" was made June 25 when pictures of a fight in New York appeared on a $20' \times 15'$ screen in Philadelphia. Every seat in the theatre "was transformed into a ringside seat at New York's Yankee Stadium."

Travels with a camera. Geography begins at home. Amherst's students in geology get such a beginning in the new teaching technique that turns a sequence of pictures into a travelogue, showing first the near-by countryside for easy comparison of pictures and original.

A bird's-eye view is hard to get by squinting at two aerial pictures. Fairchild has a stereoscope that folds like spectacles and weighs only 7½ ounces. In use by photogrammeters (or anybody who wants to get a stereoscopic view from aerial photographs), the device stands on three legs, 4 inches above the prints.

Glass. Eastman Kodak has reported a new glass formula, described by the company as a major optical development of the past fifty years. The formula includes lanthanum, tungsten, and tantalum. With this high-index glass, designers have produced lenses that give sharpness and clarity, particularly in the corners of the field.

Air Force photographers have a new telephoto lens by Eastman. This "giant" is nearly 4 feet long and 15 inches in diameter, to take $9'' \times 18''$ pictures from ten miles up. To protect against the cold of the high altitude, the mounting has a thermostatically controlled electric heater.

Quite the opposite is the Photographic Arts Manufacturing Corporation's Britar telephoto lens for Keica or Kardon. It weighs only 7½ ounces.

Speed. Last year General Electric and Eastman worked together to develop an industrial camera that will picture the millionth-second surge shown on a cathode-ray oscillograph. After the snap the camera develops the picture and projects it on a viewing screen.

Another General Electric development is apparatus that produces a brilliant flash by discharge of electricity through argon to expose objects in the field for only two millionths of a second. The camera waits, shutter open. A shell moving at more than twice the speed of sound covers only 65 thousandths of an inch in that interval.

Camera's camera. Precision makes a new Bausch & Lomb camera suitable for testing new lens formulas. The rigid frame keeps film plane and lens board parallel within .002 inch. The negatives are compared with test charts for check of resolving power, astigmatism, etc. The camera takes $5'' \times 7''$ plates, and is also used for precise photographic and copy work.

Easy reading. Microfilm use, regardless of advantages, has its bad side. It removes some of the reason for visiting distant libraries and it condemns readers to stuffy booths. American Optical Company's new microfilm reader looks comfortable (in the picture) and seems to have advantages. It weighs 70 pounds, and can be wheeled about quickly and silently. The image has no "hot spot," and the light can be adjusted to suit and save the reader's eyes. The motor has high and low speeds each way to reach a "page" quickly, and a 360° reverser to "turn" the page around.

Serious work. The "Classic 35" by Craftsmen's Guild, Hollywood, takes 35-mm pictures and yet is only $1\frac{3}{8}'' \times 2\frac{3}{8}'' \times 3\frac{3}{8}''$ over-all, "world's smallest" camera for such size film. Fuss and bother are nil, according to the makers, even to preventing double exposure or no exposure: yet the camera is efficient for "really serious work."

Drying. Photography and other processes, including some preparations in medicine, may benefit from the "Anhydrator" of Oscar Fisher Co., New York. Clean air is recirculated for drying films and materials without heat.

M. W.

COMMENTS AND CRITICISMS

CAN INTELLIGENCE BE MEASURED?

(A Reply to Davis and Havighurst)

It would appear that the crux of the article by Davis and Havighurst on "The Measurement of Mental Systems" (SM, April 1948) is the assertion that a so-called intelligence test is invalid if it gives higher average scores to the highest socioeconomic group than to the lowest socioeconomic group, on the ground that actually there is no difference between these two groups in intelligence.

This belief by the authors seems to be made very clear by the statement (p. 312, col. 2) that "any differences between the average response of different cultural groups to a mental problem may be attributed to their unlike cultures. Therefore, all problems that show socioeconomic differences in performance should be ruled out of the tests as unfair."

The assumption that all socioeconomic groups are equal in innate intelligence and that they make different average scores in intelligence tests only because of faults in the tests seems to run through the article from the first page. Yet this assumption appears to be wholly gratuitous. In the whole 16-page article there appears to be no proof whatever of this statement. The authors seem merely to give any other contention a wave of the hand with the assertion (p. 312, col. 1): "There is no evidence, and no theory shared by the leading human geneticists, to the effect that the underprivileged socioeconomic or racial groups are genetically inferior to the more prosperous socioeconomic groups." One is led to wonder who these "leading human geneticists" are. No names are given, nor is any reference given to any research or published opinion.

The authors go on to say: "There is abundant scientific evidence, advanced by sociologists and social anthropologists during the past twenty-five years, that a child's particular social and cultural environment directs, trains, and motivates his behavior." The odd fact is that nobody would contradict that statement, but it proves nothing with reference to the preceding statement. We all know that environment governs behavior, and it would seem that most scholars believe that heredity also helps govern behavior. The age-old controversy has been merely: Which has the more influence, heredity or environment? It is obvious, then, that to offer the assertion that scholars believe that environment governs behavior as proving that heredity does not govern behavior is fatuous and beneath the intelligence of the two authors.

Let us do a little logical reasoning in this matter. We all agree that there are hereditary differences in intelligence—that is, in the ability to solve life's problems and adjust oneself to the environment. If two young men have a hereditary difference in intelligence—the ability, as we said, to solve life's problems and adjust themselves to their environment—it is obvious that the favored one will be more likely to succeed socially and economically than the less favored one. This means that in the long run the less favored individuals—those with less hereditary ability to solve life's problems and adjust themselves, especially in these days of insecurity—must necessarily tend to gravitate into the lower socioeconomic groups. This means that there comes to be a tendency for low socioeconomic groups to be of lesser innate intelligence than high socioeconomic groups.

It certainly has been proved beyond question that intelligence is inherited; that is, on the whole, the children of parents of lesser intelligence tend to be themselves of lesser intelligence. This fact, coupled with the fact that parents in the lower socioeconomic groups tend to be of lesser intelligence, means that *children* in the lower socioeconomic groups tend to be of lesser intelligence.

We see, therefore, that logical reasoning brings us to the inescapable conclusion that children in the lowest socioeconomic group must be expected to make lower average scores on a valid intelligence test than those in the highest socioeconomic group.

It is interesting that the authors consider the Otis Alpha Non Verbal Test to be the most valid in that it differentiates least between socioeconomic groups. Actually, however, there would seem to be no way to determine on a priori grounds how much difference between socioeconomic groups is to be expected in average scores or in percentages of passes of items in a test; hence, it would seem that the authors' Figure 1, showing various amounts of average difference between percentage of passes for items of several tests, is of no significance in this connection. We don't know whether the test that showed small differences between the two socioeconomic groups is a better or poorer test of intelligence than the others.

One important consideration that the authors failed to take into account is that intelligence tests have various purposes, and that different kinds of validation of items are needed for the different purposes. Thus, one purpose of an intelligence test might be to study scientifically the hereditary differences in intelligence between races or socioeconomic groups, whereas another purpose might be to determine for practical purposes the probable rate at which different pupils in a given school will progress in that school.

If our purpose is of the first kind, we may need to go to great trouble and expense to get a test that is as free as possible from any influence of school environment. It might be worth while to spend five hours testing each individual with a great variety of items, including, say, tests of manual dexterity, learning ability in an actual learning situation, etc.

On the other hand, if our purpose is merely to prognosticate pupils' ability to progress through a given school and perhaps to classify them accordingly within that school, we can use a much more convenient test. This is because even the differences among pupils that may be the result of different "cultures" have a marked effect in determining pupils' rate of progress through a given school, and consequently it is of no great concern to the test maker if the score is affected to some extent by "cultural influences" so long as it makes an accurate prediction of what it is desired to predict.

The authors criticize the Otis tests because they were validated on the principle that that item is best which best differentiates between pupils who are making slow progress and those making rapid progress. But it would seem obvious that if predicting rate of progress is the chief purpose of the test, then rate of progress of pupils is the

UNIVERSITY OF MICHIGAN LIBRARIES

st criterion of validity of the items. To be sure, we not prevent cultural influences from helping to deter- ne pupils' scores but, so long as those influences also ect rate of progress in school, we need not consider m altogether vicious in influencing the pupils' scores.

As indicated above, the making of a practical and work- e intelligence test involves a compromise—some sacri- e of scientific validity for the sake of economy and iciency of administration. This means that for practical rposes the variety of types of items is necessarily mewhat restricted—to the distress of the authors.

If the authors wish merely to point out the shortcomings the present tests for such scientific uses as studying ferences in intelligence between various racial or socio- nomic groups, well and good. But in the present study authors have made the following errors:

1. It is assumed almost gratuitously that all socio- nomic groups are equal in inherent intelligence. No idence is given in support of this assertion.
2. The obvious error of such a hypothesis, as becomes ar through a little logical reasoning, is ignored.
3. The conclusion drawn from the erroneous hypothesis, mely, that any test that differentiates between high and socioeconomic groups is invalid, is necessarily er- eous.
4. No account is taken of the different uses to which elligence tests are put and of the difference in mode validation needed for the best results in the different s.
5. Tests are condemned for having been validated in manner actually best calculated to make them most ful for the purpose for which they will be used in ety-nine cases out of a hundred, that is, predicting ative rates of progress of the pupils of a given school ugh that school.

Incidentally, it is noted that the title of the article uses term "mental systems," yet the article would seem to e no explanation as to what is meant by mental systems any justification for the assumption that there are such gs as mental systems.

ARTHUR S. OTIS

shington, D. C.

(A Rejoinder to Mr. Otis)

Mr. Otis makes two basic criticisms of our paper:

First, he says that we assume that all socioeconomic ups are equal in innate intelligence. This is stating position a little more strongly than we would do; we ld rather say that in view of what is now known about etics and about intelligence testing, the safest assump- is that the several socioeconomic groups in the United tes are equal in innate intelligence. We cannot prove assertion on the basis of evidence at present available, neither can the assertion that socioeconomic groups er in innate intelligence be proved with evidence now and. (See the mathematical papers of J. B. S. Haldane, celot Hogben, Julian Huxley, and Levy on genetic ction; the studies of twins by Newman, Freeman, and zinger; and the papers of B. Kuppusawny.)

In particular, the "logical reasoning" Mr. Otis offers on matter of differences of innate intelligence between oeconomic groups leaves us unconvinced. We agree two young men may have hereditary differences in "ability to solve life's problems and adjust themselves their environment." We suggest, in passing, that in- gence tests probably do not measure this ability very , nor are they intended to do so. The argument ad- vanced by Mr. Otis is that the more innately intelligent young man, and others like him, will be more likely to succeed socially and economically than the less gifted ones. In the long run, he argues, the less intelligent ones (genetically) will gravitate into the lower socioeconomic groups, and will pass on their less favorable genetic characteristics to their descendants.

We cannot agree with this argument, on the basis of what is now known about genetics. Intelligence, in so far as it is passed on genetically, is transmitted as a highly complex set of mental characteristics on an unknown but certainly large number of widely separated genes. The kind of argument advanced by Mr. Otis might apply to such a relatively simple characteristic as eye color, if eye color had any function in socioeconomic competition, but it probably does not apply to such a complex set of characteristics as intelligence, particularly when intelligence is defined as ability to solve life's problems and adjust oneself to the environment. From what is known about genetics, the children of a man who was well favored with innate intelligence would have very little chance of being better favored than the children of a man who was less well endowed genetically in these respects. Both men carry many latent characteristics as well as manifest ones in their genetic structure. Both men, furthermore, have wives who in each conception contribute half the genes to the offspring. There is very little chance that socioeconomic differences have been translated into genetic differences of intelligence under these conditions in the course of the half a dozen generations that span the life of American society. Possibly a geneticist, with present-day knowledge and power to control human mating, could breed groups of humans with different average innate intelligence in the course of fifty or a hundred generations, but that remote possibility does not bear on the present situation.

Thus we believe the safest assumption that can be made is that there are no innate differences of intelligence between socioeconomic groups in the United States today.

But suppose we grant the possibility that there is a small innate difference of this sort (which we must consider as a possibility, though insisting that it is certainly small if it exists). Even in this case, we should point out that, with the great individual variability of innate intelligence in all socioeconomic groups, and with the higher birth rate of the lower socioeconomic groups, there are many more highly endowed children born in these groups than in the upper socioeconomic groups, because the latter groups make up a smaller fraction of the population. Hence, any use of intelligence tests that denies educational opportunity to children of the lower socioeconomic groups is a socially unhealthy act. This is the basic criticism we are making of current intelligence testing.

This brings us to the second criticism made by Mr. Otis, namely, that we do not take account of the different uses to which intelligence tests are put, and of the consequent differences in desirable modes of validation of the tests. On the contrary, we are very much concerned with the uses to which intelligence tests are put. Our observation of the use of intelligence tests shows us that where these tests are used in school systems on a large scale, they are used to determine what kinds of educational opportunity shall be given to boys and girls.

In the very first grade, children in most large cities are segregated into "fast" and "slow" groups, partly upon the basis of their intelligence-test scores. This practice is continued in many secondary schools. In many cities, discrimination against lower socioeconomic neighborhoods

or against ethnic groups, with respect to school equipment and curricula, is defended upon the basis of their low average test scores. Thus any cultural loading in the tests becomes a matter of urgent practical influence upon public education in this democracy.

The school authorities go to the expense of giving intelligence tests, not merely as Mr. Otis claims, to prognosticate pupils' ability to progress through school, but actually to put pupils into one or another kind of learning program, or to decide how much to expect from pupils. If, therefore, a boy of high innate learning ability but of experience inadequate to score high on a certain test gets a low score, he may be put into a class from which not much is expected, or he may be shunted into a curriculum that will prevent his entering college, or his parents may be advised that he is not "college material."

It is just because intelligence tests are used as a means of determining what educational opportunity children shall have, and of giving the best opportunities to children with the higher scores, that we insist that nothing but the best test of learning ability should be used in the schools. The best test of learning ability for use in American schools is one that does not penalize any group because of its cultural experience. We believe that such a test can be devised; we believe that it can be devised from problems which all children have had approximately equal chances to experience; and we expect that it will show as high a correlation with teachers' ratings of the same pupils as did the Stanford-Binet. The correlation of the Stanford-Binet test with "school success" was only 0.48.

ROBERT J. HAVIGHURST
W. ALLISON DAVIS

Department of Education
University of Chicago

THE SCIENTIST AS PHILOSOPHER

I wish to congratulate Dr. A. W. Lindsey on his article on "The Faith of Science" (SMO, May 1948). . . . It seems to me replete with good concepts and sound viewpoints.

Following through on the subject of the philosophies of scientists, which Dr. Lindsey and I agree in distinguishing from the philosophy of science, I should like to suggest that it is hardly necessary to concede anything like as much austerity to the scientist as Dr. Lindsey seems to assume. Is not the fact of the scientist's leaning toward science usually grounded upon a warmth of imaginative appreciation roused in him in youth by the visible content of this material universe? And is not his supposed austerity really something rather fictitious, attributed to the scient by bewildered nonscientists because they have failed orient themselves successfully regarding the vitally p sonal matter of the philosophies of their scientist friend Where such a group difference seems to occur betwe scientists in general and other educated persons in gener I hesitate to ascribe it to any higher loyalty to truth on part of scientists. Surely bankers, statesmen, and historia would be as truly averse to grounding their lives on fal cies as would scientists. Where there is a difference must be due rather to the nature of the training they ha had in the process of ascertaining truth, or to aspects their informational equipment.

Until due allowance has been made for the differen of mental baggage, there will be many failures to grasp deeper similarities in the human qualities of men of scie and their fellow-beings. . . .

If our scientist has less to say about faith to which r have recourse when they reach the limits of true comp hension, and more to say about an agnosticism with craving for an explanation when accurate (or reliably proximate) explanation is obviously beyond reach, also is not so much through austerity as it is throug different training in the processes of ascertaining truth different familiarity with the great cosmic reliabilities, a better appreciation of the congenial livability of unperfected, growing, solidly verified system of truths

ADDISON GUL

Department of Bio-Chemistry
University of Missouri

OF ERROR, A TRACE

It is mystifying that Professor Ferguson, who w such an interesting article on contact lenses (SM, J 1948), should confuse optometrists with opticians, possibly this accounts for his statement on page 497 optometrists "are neither trained nor qualified to diagn eye defects or prescribe lenses, regular or contact." Any familiar with the training and experience of optometr knows that Professor Ferguson's generalization has bases in fact, for there are a great many optometrist this country well qualified to determine errors of refract and to prescribe eyeglasses. A *scientific* monthly sho not supply space for such an unjustifiable attack upo recognized profession and widespread practice.

MATTHEW LUCKI

Lighting Research Laboratory
General Electric Company
Cleveland, Ohio

UNIVERSITY OF MICHIGAN LIBRARIES